Audio - Visual Interaction in P300

[1]Aparna T.H. & [2]Sandeep Maruthy

Abstract

The present study investigated the interaction in the processing of auditory and visual stimuli in the generation of P300 and the effect of electrode site on P300. P300 was obtained for auditory alone condition, audio-visual congruent and audio-visual incongruent conditions, which in turn were between two electrode sites (C_z & Pz) for latency and amplitude. The experimental data was obtained from 20 normal hearing adults. The results showed that the mean latencies were shorter in audio-visual congruent condition compared to that in auditory mode and audio-visual incongruent conditions. The mean amplitude was higher in the audio-visual congruent condition compared to auditory and audio-visual incongruent conditions in both the electrode sites. The mean latency was shorter and amplitude was higher in Pz than that in Cz in all the three stimulus conditions suggesting that distribution of the electrical field and the dipole of the potential can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain. The P300 recordings supported earlier perceptual processing and better speech perception in auditory-visual mode compared to that in auditory mode suggesting that bimodal condition facilitates the detection of the stimulus.

Keywords:*P300, audio-visual interaction, congruent, incongruent.*

Introduction

Audiovisual modality for speech perception is a well known rehabilitative strategy in individuals with hearing impairment, where only auditory modality is not providing sufficient cues. the very logic of this strategy is that the bimodal stimulation facilitates speech perception compared to unimodal stimulation. Various behavioral studies have shown that behavioral responses to auditory (A) and visual (V) stimuli are improved when those stimuli are accompanied by a task-relevant stimulus in the other modality (Odgaard, Arieh & Marks, 2004; Lippert, Logothetis & Kayser, 2007).

Physiologically, visual influences in the auditory cortex would result from feedback projections from polysensory areas (Calvert, Campbell & Brammer, 2000; Miller & D'Esposito, 2005), particularly from the superior temporal sulcus. Several studies have shown that auditory event-related potentials (ERPs) can be altered by visual speech cues as early as the N1 (Besle, Fort, Delpuech & Giard, 2004; Mottonen, Schurmann & Sams, 2004; Wassenhove, Grant & Poeppel, 2005) that is, during the building of an auditory neural representation itself (Naatanen & Winkler, 1999). Such an enhancement however also depends on the attention. Talsma, Dotty and Woldorff (2007) reported that when attention was directed to both auditory and visual modalities simultaneously, audio-visual interaction occurred in early sensory processing. However when only one modality was attended, the interaction process was delayed to later in processing. Cross-modal enhancement of neural activity has been demonstrated using electrophysiological recordings (Giard & Peronnet, 1999; Foxe, et al., 2000) as well as neuro imaging measures (Calvert, et al., 2000).

Although it is clear that audio-visual integration enhances speech perception, the neuro-physiological basis of it is yet to be completely explored. The influences of multimodal stimulation on the neural processing would throw light on the time course of multisensory interaction. It would lead to understanding of whether multisensory stimulation results in speeding up the perception process, strengthen the perception or both.

The P300 being the first true event related potential, is expected to show changes with polysensory stimulation, if any. P300 is not a modality specific response and therefore can be elicited with auditory, visual as well as somatosensory stimuli. If P300 is found useful in assessing audiovisual interaction, it can be used as a valuable tool in assessing individuals with deficits in audiovisual interaction like in dyslexia, central auditory processing disorder, autism spectrum disorder and schizophrenia. Thus the present study aimed to document the characteristics of P300 in different audiovisual conditions.

The primary objective was to investigate the effects of audiovisual interaction on the latency and amplitude of P300. The secondary objective was to study the latency and amplitude differences among Cz and Pz electrode sites, if any.

Method

The study was based on the hypothesis that cross modal interaction of auditory and visual domain has an effect on the latency and amplitude of P300. The following method was adopted to test the hypothesis.

[1]Email: aputh10@gmail.com,
[2]Reader in Audiology, Email: msandeepa@gmail.com

Participants

25 years participated in the study. All the participants had puretone thresholds within 15dB HL at octave frequencies between 250 and 8000 Hz. Their speech identification scores (SIS) were normal in both quiet and in noise (SIS above 90% in quiet & above 60% in noise). All the participants had normal middle ear functioning indicated by type A tympanogram and presence of acoustic reflexes bilaterally. They had normal or corrected-to-normal vision verified using a Snell's chart at a distance of 6 feet. They did not have any past/present history of neurological problems.

The participants were screened for central auditory processing disorders using Speech in noise test (SPIN) in which they obtained a score of >60% in both ears. All the participants were meritorious students from different parts of the country pursuing their bachelor's or master's degree in Speech and Hearing.. They were blinded to the purpose of the study and a written consent was taken from each participant prior to testing. The method used in the study was approved by the institutional review board.

Stimuli and Test Environment

Auditory as well as visual stimuli were used for recording P300. The stimuli consisted of two CVC words; /cat/ and /bat/. These two CVC words were chosen as they were minimal pairs and were picturable. A minimal pair was preferred as the two words of the pairs would elicit similar LLRs, which in turn was necessary for better P300 detectability. The pictures of the cat and bat were used as visual stimuli. The words were spoken by an adult native Kannada speaker and were recorded by a unidirectional microphone in a sound treated room. Stimuli were digitally recorded using Praat Software (version 5.1.31) at a sampling frequency of 44,100 Hz and 16 bit digitization. The stimuli were then edited for noise and hiss reduction using Adobe Audition (Version 1.5). They were normalized to the same scaling factor and were edited to restrict the duration to 250 ms.

The two auditory stimuli were presented in Odd ball stimulus paradigm. Stimulus /cat/ was presented frequently while the stimulus /bat/ was presented infrequently. In the audio-visual mode of presentation, visual stimuli which included pictures of /cat/ and /bat/ were presented only with infrequent auditory stimulus-/bat/. Depending on the visual stimuli presented, audio-visual mode was either termed as congruent or incongruent. If picture of /bat/ was presented along with the auditory infrequent /bat/, it was AV- congruent condition. On the other hand, if picture of /cat/ was presented along with the auditory infrequent /bat/, it was AV-incongruent condition. The oddball paradigm involved the presentation of one infrequent after two frequent stimuli.

The auditory stimulus was 250 ms in duration while the visual stimulus was 1000 ms in duration. The visual stimulus was triggered along with the auditory stimuli through an external laptop during audio-visual presentation. A laptop computer was used for visual stimulus presentation. Recording of the stimulus and all the evaluations were performed in sound treated rooms where the noise levels were within permissible limits (ANSI.S3.1, 1991). The rooms were also electrically insulated.

Recording of P300

To record P300, the participants were made to sit in a comfortable reclining chair and were asked to relax. The Cz and Pz electrode sites were cleaned with skin preparation gel and the disc electrodes were placed using a conduction paste. Prior to recording P300, an absolute impedance of less than 5 kOhms and relative impedance of less than 2 kOhms was ensured. The auditory stimuli were presented through insertphones at 80 dBnHL. The repetition rate used for auditory stimulus was 1/s and 1/3s for visual stimulus. The resultant auditory evoked potentials were recorded from two scalp electrode sites (Cz & Pz) referenced to the nose tip. The recorded responses were then amplified (50,000 times) and band pass filtered between 1 and 30 Hz. The responses were averaged totally for 300 stimuli and were analyzed for 750 ms.

The participants were instructed to minimize eye blinks to reduce the contamination of the desired EEG. They were instructed to pay attention to the blocks of stimuli which were presented and were asked to mentally count the infrequent stimulus (bat) during the auditory presentation mode. During audio-visual presentation mode, the participants were asked to press an arrow key in a computer keyboard, after hearing each stimulus and to watch the monitor kept in front of participant which displayed the pictures. The visual stimulus would arrive along with infrequent auditory stimulus, and was triggered by the arrow key pressed for the second frequent stimulus. The visual stimulus, although triggered by the second frequent, was displayed along with infrequent stimulus only. P300 was recorded for 3 different stimulus paradigms: The 3 stimulus paradigm included, auditory, audio-visual congruent and audio-visual incongruent. The order of the 3 paradigms was randomly used.

Analyses

The P300 was identified in each participant, for each stimulus paradigm, and at each electrode site. The response was analyzed to note down onset latency, peak latency, offset latency and the peak amplitude. The responses were subjectively analyzed by two experienced audiologists and the average waves recorded for the frequent and infrequent stimuli were compared. P300 was detected in the infrequent wave. The criteria for on-

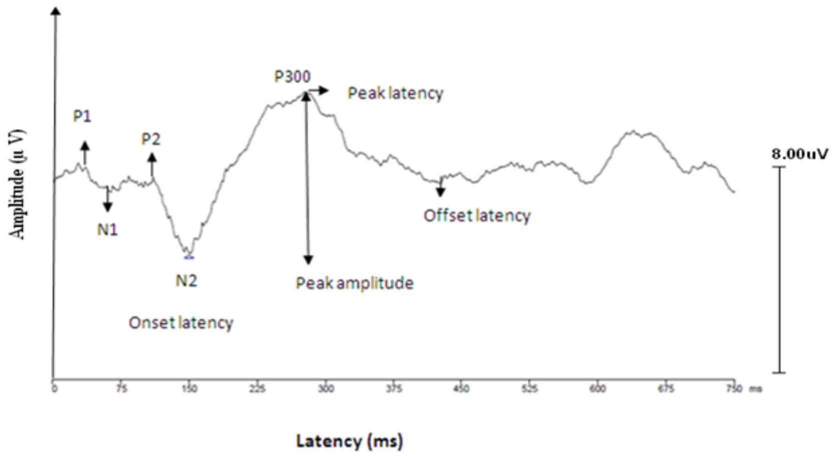


Figure 1: A representative wave recorded in auditory modality with the target parameters marked.

set and offset latency were set based on grand averaged waveforms of infrequent stimulus in each modality. A representative recording with target parameters marked is shown in Figure 1.

Results

The primary aim of the present study was to analyze the effects of audio-visual interaction on P300 latency and amplitude. The secondary aim was to evaluate the effect of electrode site on P300 latency and amplitude. While statistically testing these effects, presentation mode and electrode site were treated as independent variables, while amplitude and latency of P300 served as dependent variables. The statistical analysis was carried out using Statistical Package for Social Sciences (SPSS version 16).

Effect of Electrode site on P300

The effect of electrode site was tested for both latency and amplitudes of P300. The results are reported separately for the 2 parameters. There were 3 latency parameters measured in the P300 waves: onset latency, peak latency and offset latency. The effect of electrode site was analyzed for all the 3 parameters. Figure 2 shows the mean and standard deviation of the onset latency, peak latency and the offset latency recorded at the Cz and Pz electrode sites.

The data showed that the mean latencies at Cz were prolonged compared to that at Pz. This was true in all the three modalities and for all the three latency parameters. The observed mean differences were then tested for statistical significance on paired t-test. The results of paired t-test (Table 1) showed that there was a significant difference between the latencies recorded from the 2 electrode sites. This was true in all the three stimulus conditions. Figure 3 represents the waveforms for three stimulus conditions across the two electrode sites.

The mean and standard deviation of the peak amplitude of P300 are given in Figure 4. The mean amplitudes in Pz were higher than that in Cz in all the three stimulus conditions. The comparison of mean amplitudes between the two electrode sites was done on paired t-test separately for each stimulus conditions. Results (Table 2) showed significant difference between the two electrode sites in their mean amplitudes, in all the three stimulus conditions.

Effect of Stimulus Condition on Latency and Amplitude of P300

The mean and standard deviation of onset latency, peak latency and offset latency in the three stimulus conditions are shown in Figure 2. The effect of stimulus condition on the latency parameters was analyzed separately for the data of Cz and Pz electrode sites. In general the mean latencies were shorter in audio-visual congruent condition compared to that in auditory

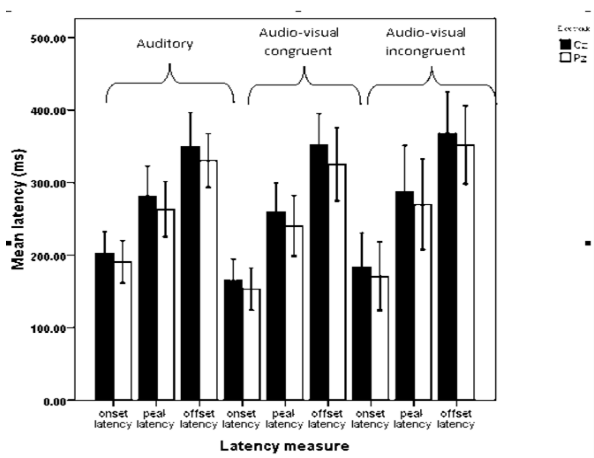


Figure 2: Mean and standard deviations of onset latency, peak latency and offset latency of P300 recorded at Cz and Pz electrode sites.

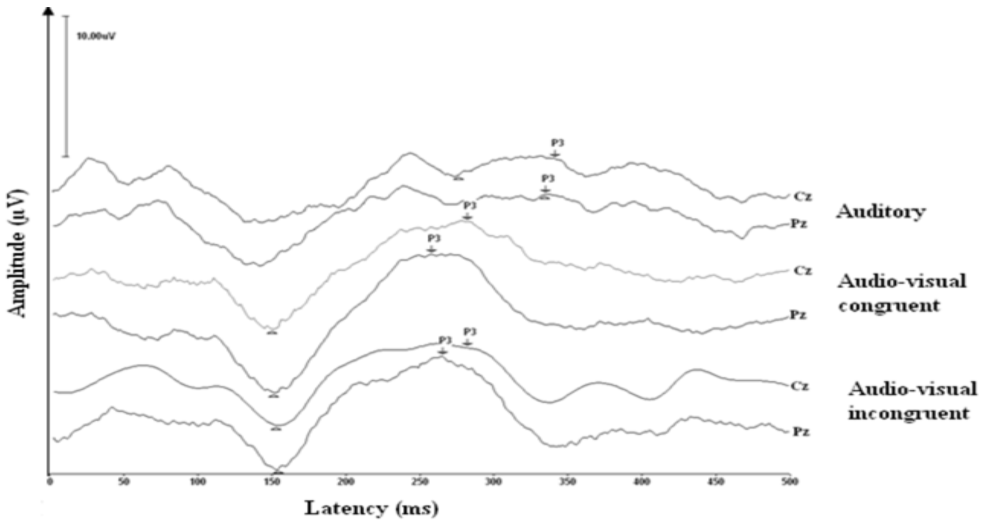


Figure 3: Waveforms for three stimulus conditions at two electrode sites.

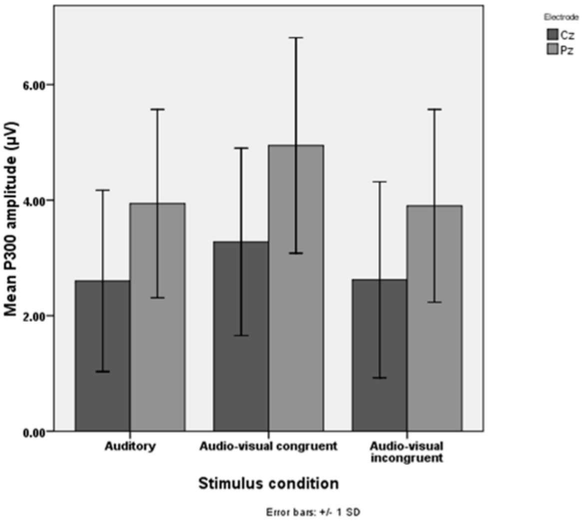


Figure 4: Mean and standard deviation of the peak amplitude of P300 in the 3 stimulus conditions.

Table 1: Results of paired t-test comparing the onset latency, peak latency and offset latency recorded from Cz and Pz electrode sites in the three stimulus conditions

Conditions	Parameters	t value	df
Auditory	Onset latency	6.222*	19
	Peak latency	1.823*	19
	Offset latency	2.618*	19
Auditory-visual (Congruent)	Onset latency	4.062*	19
	Peak latency	2.779*	19
	Offset latency	3.544*	19
Auditory-visual (Incongruent)	Onset latency	3.237*	19
	Peak latency	4.444*	19
	Offset latency	4.309*	19

*Note: *- $p<0.05$*

and audio-visual incongruent conditions. The effect of stimulus condition was tested on Repeated measures ANOVA and the results (Table 3) of Cz site showed a significant main effect of stimulus condition on only onset latency. The effect was absent on peak latency and offset latency. On the other hand at the Pz electrode site, there was significant main effect of condition on all the three latency parameters.

Wherever there was a significant main effect of condition, pair-wise comparison was tested using Bonferroni post hoc test. The results of the Bonferroni post hoc test showed that, for the data of Cz electrode site there was a significant difference between the onset latency recorded in the auditory and the audio-visual congruent condition. However there was no significant difference in the onset latency between auditory and audio-visual incongruent and, audio-visual congruent and audio-visual incongruent condition.

Similar pattern of results as in Cz electrode site was obtained at Pz site with a significant difference between the onset latency, peak latency and offset latency recorded in the auditory and the audio-visual congruent condition. However there was no significant difference in the onset latency and peak latency between auditory and audio-visual incongruent and audio-visual congruent and audio-visual incongruent conditions.

The mean and standard deviation of the amplitude recorded in the three stimulus conditions at the two electrode sites are shown in Figure 3. The mean amplitude was higher in the audio-visual congruent condition compared to auditory and audio-visual incongruent conditions. The mean data was similar in the auditory and audio-visual incongruent condition. The same trend was observed at both the electrode sites. Repeated measures ANOVA was used to test the effect of stimulus condition on amplitude and the results showed a signif-

Table 2: Results of paired t-test comparing amplitude recorded from Cz and Pz electrode sites in the three stimulus conditions

Conditions	t value	df
Auditory	-10.75*	19
Audio-visual congruent	-6.523*	19
Audio-visual incongruent	-7.556*	19

Table 3: Results of Repeated measures ANOVA for latencies across three stimulus conditions at two electrode sites

Electrode site	Parameters	F	df (error)
Cz	Onset latency	11.32*	2 (18)
	Peak latency	2.456	2 (18)
	Offset latency	1.227	2 (18)
Pz	Onset latency	11.872*	2 (18)
	Peak latency	3.487*	2 (18)
	Offset latency	3.716*	2 (18)

*Note: *- p<0.05*

Table 4: Correlation coefficient showing the relationship between index I and II derived for onset latency, peak latency, offset latency and peak amplitude

Electrode Site	Parameters	r
Cz	Onset latency	0.474*
	Peak latency	0.301
	Offset latency	0.517*
	Amplitude	0.809**
Pz	Onset latency	0.261
	Peak latency	0.481*
	Offset latency	0.444
	Amplitude	0.880**

*Note: *- p<0.05, **- p¡0.01*

icant main effect of stimulus condition on amplitude of P300 at Cz [F(2,18) = 4.291, p = 0.021] and Pz [F (2,18) = 7.208, p = 0.002].

As there was a main effect of stimulus condition, the pair-wise comparison was tested using Bonferroni post-hoc test. The results showed that audio-visual congruent condition was significantly different from auditory as well as in audio-visual incongruent condition. However there was no significant difference in amplitude between auditory and audio-visual incongruent condition. The results were same for both Cz and Pz.

Correlation of Latency Index and Amplitude Index

Results in the previous sections showed that the effect of stimulus condition was present on both latency and amplitude of P300. However, it did not give the picture of the effects of two audio-visual conditions (congruent & incongruent) related with each other.

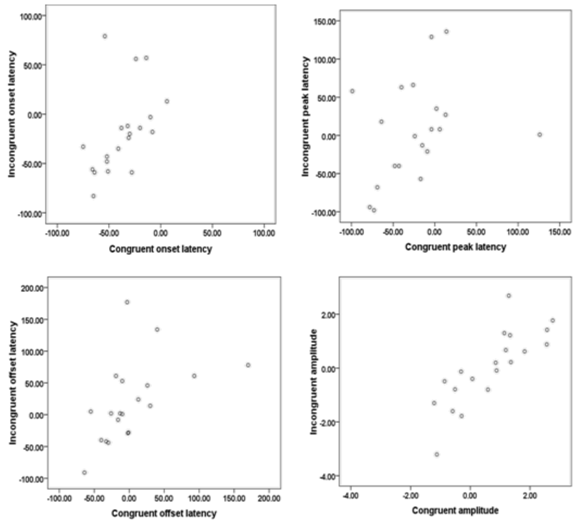


Figure 5: Scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Cz site.

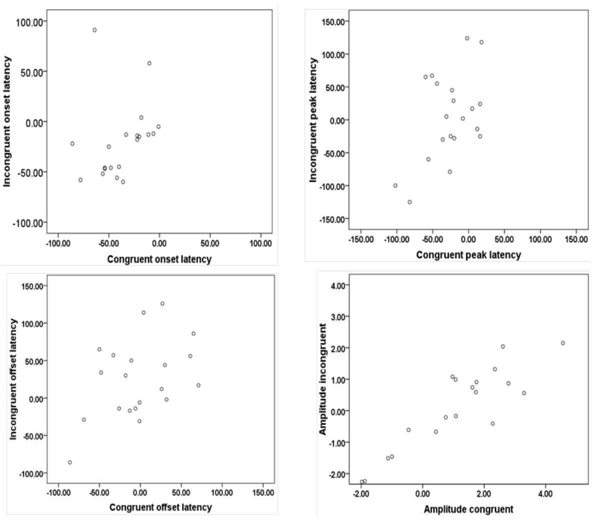


Figure 6: Scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Pz site.

To derive this relationship, latency and amplitude indexes were determined. This was done by taking the difference in the latency obtained in audio-visual conditions and auditory mode. The difference values obtained by subtracting audio-visual congruent from auditory were termed, latency index I and amplitude index I. Whereas the difference values obtained by subtracting audio-visual incongruent from auditory were termed, latency index II and amplitude index II. These indexes were calculated separately for the data from Cz and Pz electrode sites. To see the relation between the changes seen due to audio-visual congruent and AV-incongruent stimulus paradigms, Index I was correlated with Index II. This was done separately for Cz and Pz electrode site.

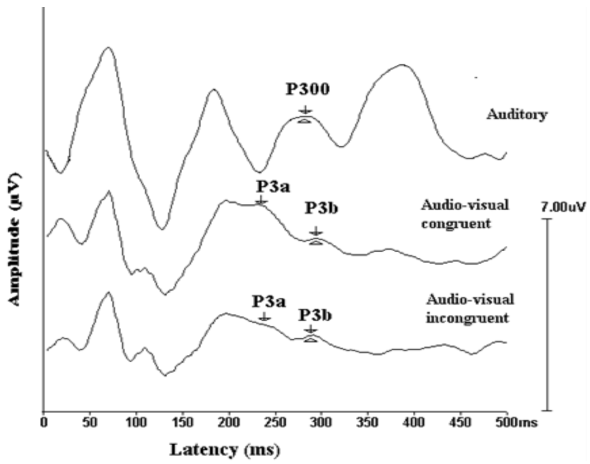


Figure 7: Mean waveforms of P300 in the auditory, audio-visual congruent and audio-visual incongruent conditions showing difference in the morphology. Auditory P300 shows single peak while audio-visual condition shows bifid peak.

Figure 5 shows scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Cz site. As can be observed in the figure, there was a direct relationship between index I and index II in all the four target parameters. The relationship was statistically tested on Pearson Correlation test. The correlation coefficient thus derived is given in Table 4. Results showed a significant low correlation between index I and index II in onset latency and offset latency. But indexes derived from the peak amplitude showed high positive correlation.

Figure 6 shows scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Pz site. As can be observed in the figure, there was a direct relationship between index I and index II in all the four target parameters. The relationship was statistically tested on Pearson Correlation test, and the results are given in Table 4. Results showed a significant low correlation between index I and index II for peak latency. But indexes derived from the peak amplitude showed high positive correlation. There was no significant correlation between index I and index II for onset latency and offset latency.

Morphological Changes in P300

For all the subjects, better waveforms were obtained at Pz compared to Cz. Comparison of P300 waveforms across auditory condition, audio-visual congruent and audio-visual incongruent conditions showed that the amplitude of P300 was larger in audio-visual conditions compared to auditory condition. It was observed that in auditory condition, a single P300 peak was seen. Whereas bifid peaks (P3a & P3b) were present in audio-visual conditions. Broadening of the P300 peak was noticed in audio-visual condition when compared to auditory condition. Figure 7 represents the mean waveforms of P300 depicting morphological changes in P300 across the three stimulus conditions.

Discussion

P300 is an endogenous potential recorded for an odd ball paradigm. Considering that the potential is generated by auditory association areas, it was hypothesized in the present study that P300 shall differ in bimodal stimulation compared to unimodal stimulation. Further, the effect of electrode site on P300 was analyzed. The analysis of independent effects of the 2 variables; mode of stimulation and the electrode site on the latency and amplitude of P300 showed some interesting findings. Attempt has been made to justify these findings and derive appropriate implications in the following sections.

Effect of Electrode Site on P300

The present study revealed that longer latencies (onset, offset & peak latency) and reduced amplitude at Cz than Pz. It is evident from this result that spatial amplitude distribution of P300 is largest at the parietal electrode sites and is attenuated as the recording site move to central locations. The exact source of the potential, region of integration and the principal components contributing cannot be inferred due to the reduced number of electrodes used in this study. However, the findings suggests that the distribution of the electrical field can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain.

These observations are in accordance with various earlier studies. Katayama and Polich (1998) obtained P300 using three stimulus odd ball paradigms and found that P300 target amplitude was larger across the midline electrode sites, being largest over the parietal site. Polich (2007) assessed correlational association between peak amplitude and latency for the P300 using auditory discrimination task. They found that for the target stimuli, P300 amplitude and latency were negatively correlated over the frontal-central and right medial/lateral recording sites. Similar results were reported by Comerchero and Polich (1999). In these studies larger P300 amplitude have been recorded at parietal sites.

Effect of stimulus conditions on latency and amplitude of P300

As evident from the results of the present study at Pz, all the 3 latency parameters (onset latency, peak latency & offset latency) were shorter in audio-visual congruent condition compared to auditory mode and audio-visual incongruent conditions. This is an electrophysiological evidence of the earlier onset of neural processing pertaining to sound discrimination, underlying

P300. Functionally, this may mean earlier perceptual processing in audio-visual congruent condition suggesting that bimodal condition accelerates the detection of the stimulus. However, this needs to be experimentally confirmed through correlation of behavioral and electrophysiological findings.

These findings are consistent with the findings of Fort, Delpuech, Pernier and Giard (2002). They assessed the effect of unimodal (auditory or visual) and non redundant bimodal conditions (auditory & visual) on ERPs. Event-related potential analysis revealed the existence of early (200 ms latency) cross modal activities in sensory specific and nonspecific cortical areas suggesting cross modal interaction resulting in facilitation effect (shorter reaction time) for the detection of bimodal stimuli compared to unimodal stimuli. Similar findings were reported in various earlier studies (Odgaard et al., 2004; Frassinetti, Bolognini & Ladavas, 2002).

Behavioral studies (Sumby & Pollack, 1954; O'Neil, 1954; Erber, 1969; Grant & Seitz, 2000., Rudmann, McCarley & Kramer, 2003; Bernstein, Auer & Takayanagi, 2004; Ross, Saint-Amour, Leavitt, Javitt & Foxe, 2007) unanimously show better speech perception in auditory-visual mode compared to that in auditory mode. The earlier latency of P300 may be the neurophysiological basis for the behavioral advantage in AV mode. Auditory-cognitive processing of information is expected to be better in auditory-visual congruent condition since the system is able to utilize the information from both the modalities in a complementary manner.

There was no significant difference in latencies (both in Cz and Pz) between auditory and audio-visual incongruent condition, also between audio-visual congruent and audio-visual incongruent condition. The exact underlying physiological reason could not be derived for this result. Logically, it was expected that the AV incongruent condition would be delayed than that in auditory mode, as the incongruency between the 2 modalities would create confusion. But the absence of the difference between the latencies of 2 stimulus conditions indicates that the processing is delayed by the incongruency. This probably is because, in instances of incongruency, the processing may be primarily determined by the dominant modality, which is auditory in this case.

The P300 picked up at Cz however showed the same trend as at Pz but only in onset latency. This could be due to differential contributions of the latent components of P300 at the 2 sites. That is, some of the latent components of P300 picked up at Pz are not picked up at Cz. Unlike at Pz, the latent components contributing for peak latency of the response at Cz were probably not sensitive to the differences in the mode of stimulation.

The present study indicated that the mean amplitude was higher in audio-visual congruent condition than auditory and audio-visual incongruent conditions. There was no significant difference in mean amplitude between auditory and audio-visual incongruent conditions. The finding again can be justified with the integration of the complimentary cues provided in the auditory and visual modes. Because the cues provided in the incongruent condition were not complimentary, enhancements in the amplitude were not seen in the audio-visual incongruent condition. Li, Wu and Touge (2010) investigated auditory detection enhancement by cross modal audio-visual interaction, by comparing the ERPs elicited by the audio-visual stimuli to the sum of the ERPs elicited by the visual and auditory stimuli. Results suggested that behavioral detection of an auditory stimulus is enhanced by the interaction of auditory and visual stimuli around 300ms in polysensory regions of the brain.

Relation between Changes Seen in Congruent and Incongruent Conditions

Because the mean changes in audio-visual incongruent condition were not significantly different from that in auditory condition, it was of interest to study the relation between the changes seen in the 2 audio-visual conditions. This would help in inferring whether the changes seen in the audio-visual incongruent were facilitative or deleterious.

Results showed a significant positive correlation between the 2 indices derived. This supports that the changes seen in the audio-visual incongruent were facilitative, like in audio-visual congruent condition, although not to a noticeable extent.

Overall, from the findings it can be derived that bimodal presentation, neurophysiologically, is beneficial over unimodal presentation for speech processing. Within bimodal conditions, congruent condition is more facilitative compared to incongruent condition and P300 can evidence the modality-based changes in the cortical auditory processing.

Conclusions

From the findings of the study it can be inferred that the distribution of the electrical field and the dipole of the potential can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain. The P300 recordings supported earlier perceptual processing and better speech processing in auditory-visual mode compared to that in auditory mode suggesting that bimodal condition accelerates the detection of the stimulus.

The P300 picked up at Cz however showed the same trend as in Pz but only in onset latency indicating that the latent components contributing for peak latency of the response at Cz were probably not sensitive to the differences in the mode of stimulation. But, in instances of incongruency, the processing may be primarily determined by the dominant modality, which is auditory in this study. Finally, it is concluded that both changes seen in audio-visual congruent and audio-visual incongruent conditions are facilitative but in audio-visual congruent condition the facilitation in evident.

The findings of the present study have important implications in diagnostic as well as rehabilitative audiology. The present study proved that P300 can be an objective index of facilitation by bimodal presentation (audio-visual). Hence it can be used in the assessment of cross modality integration at the cortical level in individuals with auditory processing disorders. P300 can also be used to monitor the progress secondary to training of cross modality integration.

References

American National Standard Institute (1991). *Specification for Maximum permissible ambient noise levels (MPANLs) allowed in an audiometric test room.* ANSI S3.1-1999 (R 2003). New York: American National Standard Institute.

Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*(1), 5-18.

Besle, J., Fort, A., Delpuech, C., & Giard, M. H. (2004). Bimodal speech: Early suppressive visual effects in the human auditory cortex. *European Journal of Neuroscience, 20,* 2225-2234.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in human heteromodal cortex. *Current Biology, 10,* 649-657.

Comerchero, M.D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110,* 24-30.

Erber, N. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research, 12*, 423-425.

Frassinetti F., Bolognini N., & Ladavas E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*(3), 332-43.

Fort, A., Delpuech, C., Pernier, J., & Giard M. (2002). Dynamics of Cortico-subcortical Cross-modal Operations Involved in Audio-visual Object Detection in Humans. *Cerebral Cortex, 12*(10), 1031-1039.

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research, 10,* 77-83.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11,* 473-490.

Grant, K. W., & Seitz, P. F. (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. *Journal of the Acoustical Society of America, 107,* 1000-1011.

Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology, 35,* 23-33.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38,* 557-577.

Li, Q., Wu, J., & Touge, T. (2010). Audio-visual interaction enhances auditory detection in late stage: an event-related potential study. *NeuroReport, 21,* 173-178.

Lippert, M., Logothetis, N. K., & Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain Research, 1173,* 102-109.

Miller, L. M., & D'esposito, M. (2005). Perceptual Fusion and Stimulus Coincidence in the Cross-Modal Integration of Speech. *The Journal of Neuroscience, 25,* 5884-5893.

Mottonen, R., Schurmann, M., & Sams, M. (2004). Time course of multisensory interactions during audio-visual speech perception in humans: a magnetoencephalographic study. *Neuroscience Letters, 363,* 112-115.

Naatanen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125,* 826-859.

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Behavioral Neuroscience, 4,* 127-132.

O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech & Hearing Disorders, 19,* 429-439.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10) 2128-2148.

Ross, L. A., Saint-Amour, D., Leavitt, V., Javitt, D. C., & Foxe J. J. (2007). Do you see what I'm saying? Optimal Visual Enhancement of Speech Comprehension in noisy environments. *Cerebral Cortex, 17*(5), 1147-53.

Rudmann, D. S., McCarley, J. S., & Kramer, A. F. (2003). Bimodal display augmentation for improved speech comprehension. *Human Factors, 45,* 329-336.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise, *The Journal of the Acoustical Society of America, 26,* 212-215.

Talsma, D., Doty, T. J., & Woldorff, M. G.(2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17,* 691-701.

Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 1181-1186.

Lexical Neighbourhood Test (LNT) for Children in Kannada

[1]Apoorva H.M. & [2]Asha Yathiraj

Abstract

The aim of the present study was to develop a Lexical Neighbourhood Test in Kannada for children aged 6 to 8 year. The study also aimed to check the utility of the developed test on the target group and in children with hearing impairment using cochlear implants. The developed test material consisted of two word lists containing 40 words each which were constructed taking into account the lexical density as well as the frequency of occurrence of the words. Each list contained 20 lexically easy words and 20 lexically hard words. The developed test material was administered on both the groups. Results revealed that there was no significant difference between younger (6 to 6; 11 years) and older (7 to 8 years) normal hearing children. There was no significant difference between the two lists developed thus confirming inter-list equivalency. Both the groups performed better on the easy words and poorer on the hard words and thus there was a significant effect of lexical difficulty on both the groups. There was a significant difference between the performances of children with normal hearing and children with hearing impairment on both the lists for easy and hard words. The results demonstrated that the developed material represented different lexical difficulties. Hence, it can be used to tap the perceptual differences in children and can serve as a valid clinical tool in examining perceptual processes underlying spoken word recognition in Kannada.

Keywords: *Lexical density, easy words, hard words, speech identification*

Introduction

As early as 1952, Hirsh et al. noted that speech tests have diagnostic and prognostic values. Speech audiometry continues to be considered essential due to its inherent advantage over pure-tone audiometry. Speech stimuli have been found to aid in detecting perceptual difficulties that may go undetected if only pure-tones were used. Pathologies in the retro-cochlear region and higher auditory pathways have been reported to not manifest themselves if evaluated only using pure-tones. This occurs despite the presence of marked speech perception difficulties. Hence, speech has been considered the preferred test material for assessing higher cortical functions (Jerger & Hayes, 1971; Jerger & Jerger, 1974). It has also been used in assessing the success in otological surgery (Kasden & Robinson, 1970), in determining communication abilities (Berger, Keating & Rose, 1971) and for hearing aid evaluation (Markides, 1977). Further, speech audiometry has been noted to help in early detection of slight losses which are otherwise overlooked (Ritchie & Merklein, 1972). The need for speech audiometry has arisen mainly because speech is by far the most important class of sound that one hears.

Several speech identification tests have been developed for children (Ross & Lerman, 1970; Mayadevi, 1974; Elliot & Katz, 1980; Moog & Geers, 1990; Rout, 1996; Vandana, 1998; Prakash, 1999; Begum, 2000; Chowdary, 2003; Jijo & Yathiraj, 2008). However, such tests have been observed not to address difficulties faced by all children, especially those using cochlear implants (Cullington, 2000) as these children greatly vary in their spoken word recognition skills (Staller, Beiter, Brimacombe, Mecklenburg & Arndt, 1991). This variability depends in part on the age of onset and duration of their hearing loss (Fryauf-Bertschy, Tyler, Kelsay & Gantz, 1992), and on the length of cochlear implant use (Fryauf-Bertschy et al., 1992; Miyamoto et al., 1992; 1994). Thus, as children develop, there is a great need to have different levels of tests that increase in terms of their perceptual difficulty.

In order to assess the developing auditory skills in children as they grow, a number of speech perception measures have been developed. These speech perception tests range from very simple closed set tests like Pattern Perception Test (Moog & Geers, 1990), Word Intelligibility by Picture Identification test (Ross & Lerman, 1970), Northwestern University-Children's Perception of Speech (Elliot & Katz, 1980), Picture speech identification test for children in Tamil (Prakash, 1999) and in Hindi (Chowdry, 2003), Early Speech Perception test for Malayalam speaking children (Jijo & Yathiraj, 2008) to more complex open-set tests such as Phonetically balanced word lists-Kindergarten (Haskins, 1949), Bamford-Knowal-Bench sentence test (Bench, Kowal & Bamford, 1979), Common phrase test (Robbins, Renshaw & Osberger, 1995) . However, it has been reported by Mukari, Ling and Ghani (2007) that children with cochlear implants performed poorly on the PB-K. This has been attributed to the test containing words that are unfamiliar to young deaf children who typically have very limited vocabulary.

The Lexical Neighbourhood Test (LNT) was developed by Kirk, Pisoni and Osberger (1995) to assess spoken word recognition in order to reveal the perceptual processes employed by children, especially for those using cochlear implants. The LNT test items were formed

[1]Email: apoorvahm88@gmail.com,
[2]Professor of Audiology, Email:asha_yathiraj@rediffmail.com

based on the frequency of occurrence of word in the language (word frequency) and the number of phonetically similar words surrounding the word (lexical density). Words with many lexical neighbours were considered to have a 'dense' lexical neighbourhood, whereas those with only a few lexical neighbours were considered to have a 'sparse' neighbourhood. Based on word frequency and lexical density, 'lexically easy' and 'lexically hard' words were categorized. The lexically 'easy' words had more frequently occurring words with 'sparse' neighbourhood and lexically 'hard' words had less frequently occurring words with 'dense' neighbourhood.

The Lexical Neighbourhood Test is considered to be very useful for measuring word recognition in children with multichannel cochlear implants who exhibit varying speech perception abilities. Thus, it is reported to provide reliable estimates of spoken word recognition abilities in these children. The test is believed to allow the examination of the perceptual processes underlying spoken word recognition. Further, the test is used to gain further knowledge about the organization of sound patterns of words in young children's lexicons and the processes used to access these patterns in traditional speech identification tests (Kirk et al., 1995).

LNT has been developed in various languages like Mandarin (Yang & Wu, 2005), Cantonese (Yuen et al., 2008), and Chinese (Liu et al., 2011). In India, it has been developed in Indian English by Patro and Yathiraj (2010) and in Hindi by Singh (2010).

Though there are several speech identification tests for children developed in India (Swarnalatha, 1972; Mayadevi, 1974; Vandana, 1998; Prakash, 1999; Chowdary, 2003; Jijo & Yathiraj, 2008), they have been developed taking into account only the familiarity of the stimuli. However, it is evident from literature (Kirk et al., 1995; Dirks, Takayana & Moshfegh, 2001; Krull, Choi, Kirk, Prusick & French, 2010) that despite words being familiar, word frequency and lexical density, also need to be considered. This would provide perceptual information of two different dimensions regarding the perception of speech. This would be especially useful in evaluating the subtle perceptual difficulties of children using listening devices such as hearing aids or cochlear implants. The results from the test would serve as guidelines regarding the usefulness of the devices. In addition, the test would help determine the progress made by children using these devices over time, with or without training. In India it is currently available only in Indian English (Patro & Yathiraj, 2010) and in Hindi (Singh, 2010).

When speech is used as stimuli for assessment, the regional language used for testing becomes an important variable (Alusi, Hinchcliffe, Ingham, Knight & North, 1974). An individual's perception of speech has been found to be influenced by his first language or mother tongue (Singh, 1966; Gat & Keith, 1978). India being a multilingual country, it is necessary to develop tests for the paediatric population in different Indian languages. Currently, LNT is not available in Kannada and hence there is a need to develop such a test in the language.

The aim of the present study was to develop a Lexical Neighbourhood test in Kannada for children aged 6-8 years. The study also aimed to check the utility of the developed test on the target group and in children with hearing impairment using cochlear implants.

Method

Participants

The study was carried out in two phases. The first phase involved the development of the Lexical Neighbourhood Test in Kannada for children aged 6 to 8 years. In the second phase, the developed material was administered on normal hearing children and children with hearing impairment using cochlear implants.

Participants

For the development of the material (Phase I of the study), 45 adults and 10 children were involved. All these participants were fluent speakers of Kannada. Among the adults, 3 were regular school teachers and 3 were special educators. The familiarity of the material was tested on the ten children who were exposed to Kannada from early childhood. These children were reported to be typically developing and had no history of speech and hearing problems. In Phase II of the study, data were collected from 30 typically developing children and 6 children with hearing impairment who used cochlear implants. All the children were exposed to Kannada from early childhood.

The typically developing children (Group I) were in the age range of 6 to 8 years and were exposed to Kannada from early childhood. Their AC and BC thresholds were within 15 dB HL in the frequencies from 250 Hz to 8000 Hz and 250 Hz to 4000 Hz respectively. Their speech identification scores were 90% or higher at 40 dB SL (ref: PTA) in both ears on the speech identification tests for Kannada speaking children (Vandana, 1998). They had bilateral A-type tympanograms with acoustic reflexes present at 90 dB HL in both ears at 500 Hz, 1000 Hz and 2000 Hz. In addition, it was ensured that they had no history of any speech, language or hearing problem as well as no otologic / neurologic problems. They had no illness on the day of testing.

The children using Nucleus cochlear implants (Group-II) were tested using their device in the prescribed settings for everyday use. Of them, 3 used CP 810 sound processor, 2 used Freedom sound processor and 1 used

Sprint processor. All the participants used their device regularly for at least one year. Their aided audiogram was within the speech spectrum. The children had a language age of at least 6 years, as evaluated using the Kannada Language Test (UNICEF funded project, 1990) with clear speech, with limited number of misarticulation.

Testing Environment

All the testing was carried in a sound-treated suite. The noise levels were maintained within permissible limits, as per ANSI S3.1- 1991.

Instrumentation:

A calibrated two channel diagnostic audiometer, ORBITER-922, version-2 coupled with headphones (TDH-39), bone vibrator (B-71) and sound-field speakers were used to estimate the pure-tone thresholds and speech identification abilities. A calibrated middle ear analyzer, GSI Tympstar version-2 was used to carry out immittance tests. A computer with Adobe Audition (version 1.5) software was utilized to record and present the speech tests.

Test Procedure

Phase I - Procedure for development of the material: The development of the material involved three steps. They included determining the familiarity of words that were considered to be in the vocabulary of 6 year old children; checking the lexical density of the familiar words; and determining the frequency of occurrence of the familiar words.

Initially, a list of words that were considered to be in the vocabulary of children aged 6 years was made. The words were selected from age appropriate print material. Additionally, 15 adults who were exposed to Kannada since early childhood were asked to make lists of words which they thought would be present in the vocabulary of children aged 6 years. The words from the print material and the 15 adults were pooled into a single list by omitting the common words. Thus, 550 words were obtained and the familiarity of these words was checked on 10 children aged 6 years. A word was considered highly familiar only if the children were able to describe its meaning. Words which could be described correctly by more than 80% of the children were included for further steps in the construction of the test. Of the 550 words, 230 words were found to be highly familiar.

The lexical density of the 230 words was calculated by determining the number of lexical neighbours for each word. This was done by determining the number of words that could be formed by adding, deleting, or substituting one phoneme at a time from the target word. To carry out this task, 12 adults who were educated in Kannada from early childhood and who spoke the language fluently were considered. They were instructed to construct as many lexical neighbours as possible for each of the 230 words by making use of the procedure to calculate the same. Later, the responses from the 12 participants were pooled into a single list by eliminating the words that were repeated by the different participants. The lexical neighbourhood density ranged from 0 to 15. Those words that had 3 and less neighbours were categorized to have a 'sparse neighbourhood' and those words which had more than 3 neighbours were considered to have a 'dense neighbourhood'. This cut-off value was selected since it approximated the value recommended by Kirk et al. (1995).

Further, the frequency of occurrence of the 230 words was determined by calculating the number of times each of the words occurred in text books / story books used by children in the age range of 6 to 8 years. The text material had as many as 507 pages and 24,980 words. The frequency of occurrence of each of the words in the text material was calculated manually. Since the word count was done manually it was calculated separately by two adults who read Kannada fluently. This was done to verify the accuracy of the word count. The frequency of occurrence was noted to range from 1 to 282. The words were then divided into two groups. Those words that occurred more than 6 times in the text material were classified as 'frequently occurring words' and those which occurred 6 or less were classified as 'infrequently occurring words'. Six was taken as the cut-off point as this value divided the words into half.

Based on the frequency of occurrence and lexical density, the words were categorized as 'lexically easy' and 'lexically hard' words. The lexically 'easy' words had 'more frequently occurring' words with 'sparse neighbourhood'. The lexically 'hard' words had 'less frequently occurring' words with 'dense neighbourhood'.

Thus, two word lists containing 40 words each were constructed taking into account the lexical density as well as the frequency of occurrence of the words. Each list consisted of equal number of 'hard' and 'easy' words, i.e., 20 easy words and 20 hard words which were randomized (Appendix).

Recording of developed word-lists: The developed word-lists were recorded by a fluent Kannada speaker. It was ensured that the fundamental frequency of the speaker was within normal limits as measured on the Vaghmi software. The recorded material was edited and scaled using Adobe Audition software to ensure that the intensity of all sounds were similar. The recording was done using a sampling rate of 44.1 kHz and 32-bit analogue-to-digital converter. A directional boom microphone, placed 6 cm from the mouth of the speaker and connected to a computer was used for the recording.

An inter-stimulus interval of 4 seconds was inserted between word pairs to obtain the response from the listeners. A 1 kHz calibration tone was inserted prior to each list. The developed material was subjected to a goodness test. This was done on 10 adults to ensure that the recording was clear. All the recorded words were found to be intelligible by 90% of these adults.

Phase II - Administration of the developed test material:

Administration on normal hearing children: The recorded stimuli were played using Adobe Audition (version 1.5) which was loaded on a computer. The output from the computer was routed to sound field speakers via a diagnostic audiometer. The loud speaker was placed at 0 ^{0}azimuth at 1 meter distance from the head of the listener. Prior to the presentation of the stimuli, the 1 kHz calibration tone was used to adjust the VU meter deflection of the audiometer to '0'. The stimuli were presented in a sound-field condition at 40 dB SL (ref PTA).

The participants were instructed to repeat what they heard. To ensure that they understood the instructions, they first listened to orally presented practice items that were not a part of the test items. Following this, they were presented the recorded test items. Half of the children were presented with list 1 first and the other half received list 2 first. This was done to eliminate the list order effect. The verbal outputs of the participants were noted by the tester on a response sheet which was later scored.

*Administration on children with hearing aids / cochlear implants:*The children with hearing impairment were tested with their cochlear implants using their preferred stable map. Children using Nucleus cochlear implants were tested using the 'everyday' pre-processing strategy. The procedure followed was similar to that used to evaluate the normal hearing children. The verbal output of the participants was audio recorded and also noted by the tester on a response sheet. The response of children who had misarticulations were scored correct if the stimulus and their verbal output corresponded with the findings of an articulation test that had earlier been administered on them.

Scoring

The responses of the participants were scored by the tester. Both word and phoneme scores were calculated. For calculating the word scores, a correctly identified word was assigned a score of 'one' and a wrong answer a score of 'zero'. Thus, the maximum word score was 40 for each list. Similarly, while calculating phoneme scores, every correctly identified phoneme was given a score of 'one' and a wrong phoneme a score of 'zero'. For list-1 the maximum phoneme score was 174 and for list-2 it was 168.

Statistical Analyses

The data obtained from the 30 typically developing children and 6 children with hearing impairment were analysed using Statistical Package for Social Sciences (SPSS) software version 18. Initially, mixed ANOVA was carried out to study the main effects and the interaction between the variables for group I (children with normal hearing). Further analysis using paired t-test or Wilcoxon signed rank test was done to study the effect of lexical difficulty (easy words vs. hard words) and effect of word lists (inter-list equivalency). The performance of the children with hearing impairment using cochlear implants was analysed using Wilcoxon signed rank test. Later, the performance of the two groups was compared and analysed using Mann-Whitney U test.

Results and Discussion

The responses of 30 normal hearing children aged 6 to 8 years and 6 children with hearing impairment using cochlear implants on the developed Lexical Neighbourhood Test in Kannada are discussed. The results of the two groups of participants are discussed separately.

Performance of Typically Developing Children (group I)

To study the main effects and the interaction between the variables 2-way repeated mixed ANOVA (2 lists x 2 word types x 2 age groups) was carried out. The results of the mixed ANOVA revealed a significant interaction between the list, word type and age [$F (1, 28) = 3.094$, $p < 0.05$] and also between the list and word type [$F (1, 28) = 0.674$, $p < 0.05$]. However, there was no significant interaction between type and age [$F (1, 28) = 1.674$, $p > 0.05$] and between list and age [$F (1, 28) = 0.916$, $p > 0.05$]. The results of the effect of age, list and word type are provided below.

Effect of Age: The effect of age was determined by comparing the word scores of the two age sub-groups (6 to 6; 11 years and 7 to 8 years) on the easy and hard words for the two lists. Table 1 gives the mean, minimum, maximum and standard deviation (SD) values for the two age groups. As can be seen from the table, the scores were slightly higher for the older age group compared to the younger group for both easy and hard words on the two lists. However, the mixed ANOVA indicated there was no significant difference between the two age groups [$F (1, 28) = 2.754$, $p > 0.05$].

The effect of age was also determined by comparing the phoneme scores of the two age sub-groups on the easy and hard words. Table 2 gives the mean, minimum, maximum and SD values for the two age groups. To see if there was difference between the two age groups for the phoneme scores, Mann Whitney U test was administered. This non-parametric test was administered since

Table 1: Mean and Standard Deviation (SD) of the word scores (easy words, hard words and the total word scores) for the two age groups

Age groups		List 1			List 2		
		*Easy	*Hard	#Total	*Easy	*Hard	#Total
6-6;11 yrs N = 15	Mean	19.27	17.33	36.6	19.47	18.27	37.73
	Min	17	15	34	18	15	35
	Max	20	20	40	20	20	40
	SD	0.88	1.45	1.96	0.74	1.34	1.67
7-8 yrs N = 15	Mean	19.40	18.47	37.87	19.73	18.53	38.27
	Min	17	15	34	19	15	34
	Max	20	20	40	20	20	40
	SD	0.83	1.46	2.03	0.46	1.36	1.58

*Note: *Maximum scores for easy and hard words = 20; #Maximum Total score = 40*

Table 2: Mean and Standard Deviation (SD) of the percentage phoneme scores (easy words, hard words and the total word scores) for the two age groups

Age groups		List 1			List 2		
		Easy	Hard	Total	Easy	Hard	Total
6-6;11 yrs N = 15	Mean	98.85	96.46	97.78	99.20	97.99	98.46
	Min	94.55	93.83	94.83	95.18	94.19	95.86
	Max	100.00	100.00	100.00	100.00	100.00	100.00
	SD	1.70	1.86	1.45	1.34	1.55	1.16
7-8 yrs N = 15	Mean	99.14	97.72	98.47	99.52	98.30	98.90
	Min	96.77	92.89	95.98	97.59	94.19	96.45
	Max	100.00	100.00	100.00	100.00	100.00	100.00
	SD	1.09	2.33	1.51	0.76	1.58	1.00

the scores had to be converted to percentage scores due to the unequal raw phoneme scores across the two lists / word-type. The analysis revealed that there was no significant difference for phoneme scores between the two age groups for easy words ($Z = 0.269$, $p > 0.05$), hard words ($Z = 1.788$, $p > 0.05$) as well as total scores ($Z = 1.101$, $p > 0.05$).

The finding of the present study regarding the performance of the two age groups was similar to that of Liu et al. (2011). They studied the effect of age using the standardized Chinese version of LNT on children aged 5.0 to 5;10 years and 6.0 to 6;11. The finding of the present study and that of Liu et al. indicated that speech identification abilities of children for stimuli such as the LNT reached its peak value by a younger age. Due to a ceiling effect, with increase in age there was no further improvement in speech recognition. This probably occurred in the both the studies since the test stimuli were derived from speech materials meant for children in the younger age. While in the present study the stimuli were derived from material for the younger age group (6 to 6;11 years), in the study by Liu et al. it was derived from children younger (3 to 5 years) than their youngest age group (5.0 to 5;10 years). This could have lead to the lack of statistically significant differences between the age groups. Thus, for the Kannada LNT, children as young as 6 years can be expected to perform similar to older children.

Table 3: Mean and SD of word scores (easy words, hard words and the total word scores) for 30 normal hearing individuals

List	Word type	Mean	SD	Range Min	Range Max
List 1	*Easy	19.33	0.84	17	20
	*Hard	17.90	1.54	15	20
	#Total	37.23	2.06	34	40
List 2	*Easy	19.60	0.62	18	20
	*Hard	18.40	1.33	15	20
	#Total	38.00	1.62	34	40

*Note: *Maximum scores for easy and hard words = 20; #Maximum Total score = 40*

Table 4: Mean and SD of percentage phoneme scores (easy words, hard words and the total word scores) for 30 normal hearing individuals

List	Word type	Mean	SD	Range Min	Range Max
List 1	Easy	99.00	1.41	93.55	100
	Hard	97.09	1.67	92.89	100
	Total	98.12	1.50	94.83	100
List 2	Easy	99.36	1.08	95.18	100
	Hard	98.14	1.55	94.19	100
	Total	98.68	1.09	95.86	100

Table 5: Mean and Standard Deviation (SD) of the word scores (easy words, hard words and the total word scores) for 6 children with hearing impairment

List	Word type	Mean	SD	Range Min	Range Max
List 1	*Easy	15.83	0.98	14	17
	*Hard	12.50	1.51	10	14
	#Total	28.33	2.42	24	31
List 2	*Easy	15.83	1.17	14	17
	*Hard	12.33	1.63	11	15
	#Total	28.17	2.48	25	32

*Note: *Maximum scores for easy and hard words = 20; #Maximum Total score = 40*

Table 6: Mean and Standard Deviation (SD) of the percentage phoneme scores (easy words, hard words and the total word scores) for 6 children with hearing impairment

List	Word type	Mean	SD	Range Min	Range Max
List 1	Easy	90.32	3.85	83.87	94.62
	Hard	87.65	4.06	80.25	91.36
	Total	89.08	3.86	82.18	93.10
List 2	Easy	91.57	2.80	86.75	93.98
	Hard	87.02	3.07	83.74	90.70
	Total	88.76	2.97	85.21	92.30

*Note: *Maximum scores for easy and hard words = 20; #Maximum Total score = 40*

Since the mixed ANOVA results for word scores and Mann Whitney U test for phoneme scores revealed no significant age effect, the data obtained for the two age groups were combined for further analysis. This was done while determining the effect of inter-list equivalency and lexical difficulty.

Inter-list equivalency: To compare the performance of the normal hearing children across the two lists developed, the mean and the SD of both easy and hard words were analysed. In both list 1 and list 2, the mean word scores were similar for the easy words (19.33 & 19.60 respectively), hard words (17.90 & 18.40 respectively) and total scores (37.23 & 38.00 respectively) as evident in Table 3. Similar to the word scores, the mean phoneme scores were similar for the easy words (99.00 & 99.36 respectively), hard words (97.09 & 98.14 respectively) and total scores (98.12 & 98.68 respectively) as seen in Table 4.

Mixed ANOVA results revealed that there was no significant main effect for list [$F (1, 28) = 4.003$, $p > 0.05$]. Further, to confirm if the two lists were equivalent statistically for easy words, hard words and for total word scores, paired-t test was carried out. This analysis also revealed that there was no significant difference between the two lists for easy words ($t = 1.161$, $p > 0.05$), hard words ($t = 1.675$, $p > 0.05$) and also for the total scores ($t = 2.004$, $p > 0.05$). Similarly, to confirm list equivalency for the phoneme scores, Wilcoxon Signed Rank test was carried out. Once again, the results revealed that there was no significant difference between two lists for the easy words ($Z = 1.085$, $p > 0.05$), hard words ($Z = 2.585$, $p > 0.05$) as well as the total scores ($Z = 1.590$, $p > 0.05$).

As the two lists were found to be equivalent, it is recommended that they can be used alternatively during any assessment in order to avoid any word familiarity effect. The lists can be used interchangeably irrespective of whether word or phoneme scores are being calculated.

Effect of lexical difficulty on speech identification scores: The effect of lexical difficulty was analyzed by comparing the scores obtained for lexically easy and lexically hard words. From the mean and SD values of the word scores for easy and hard words (Table 3) it is evident that the mean scores for the latter were lesser than that of the former. Mixed ANOVA results revealed a significant main effect for word type [$F (1, 28) = 46.43$, $p < 0.001$] with lists combined. To check the effect of lexical density for each of the lists, paired-t test was done. It was found that there was a significant difference between easy and hard words for both list 1 ($t = 5.68$, $p < 0.001$) and list 2 ($t = 5.07$, $p < 0.001$).

Similarly, Table 4 shows the mean and SD values of the phoneme scores for easy and hard words for the two lists. It can be seen from the table that the mean scores for hard words were lesser than that of easy words. Wilcoxon signed rank test revealed that there was a statistically significant difference between the scores obtained for the easy and hard words. This was evident in both list 1 ($Z = 3.69$, $p < 0.001$) and in list 2 ($Z = 3.18$, $p < 0.001$).

Earlier reported studies had also observed that easy words were easier to perceive than hard words. Thus, the finding of the current study are consistent with what has been found in the English LNT (Kirk et al., 1995; Dirks et al., 2001) and in LNT of other languages (Yang

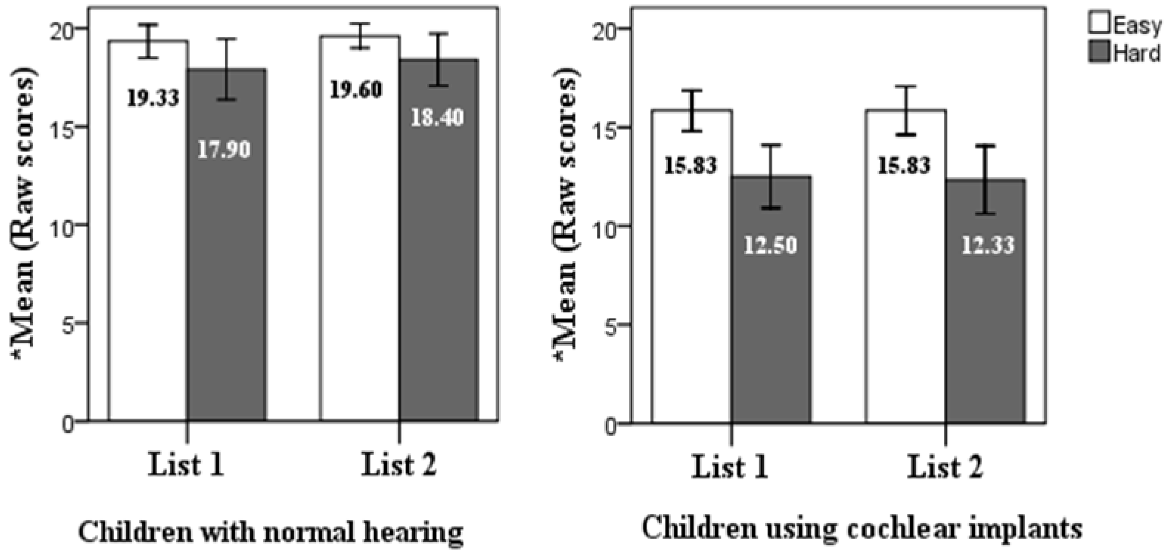


Figure 1: Mean easy and hard word scores for children with normal hearing and for children with hearing impairment using cochlear implants.

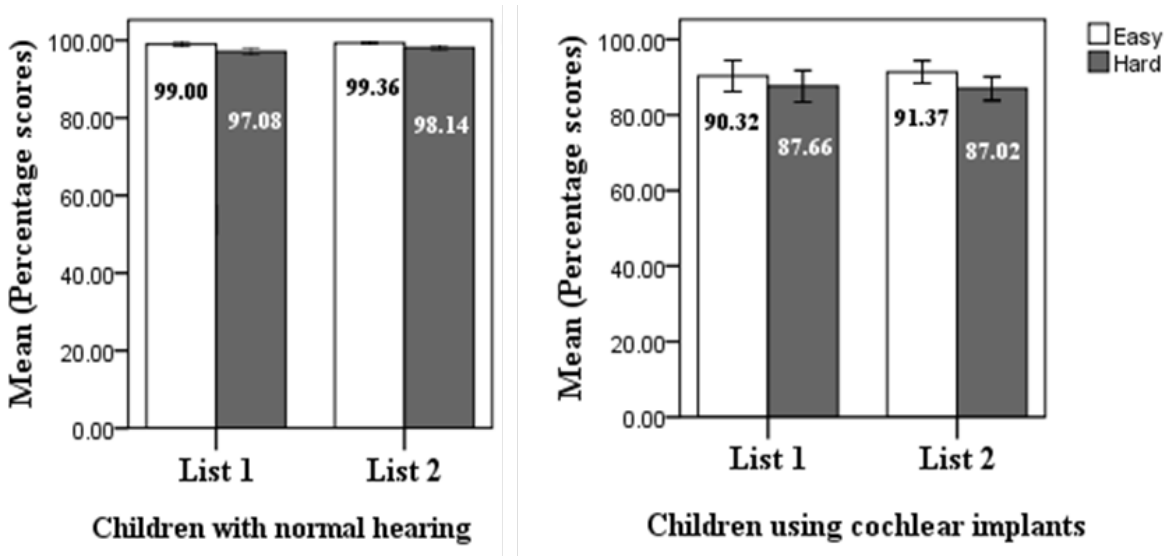


Figure 2: Mean percentage phoneme scores (for easy and hard words) for children with normal hearing and for children with hearing impairment using cochlear implants.

& Wu, 2005; Patro & Yathiraj, 2010; Singh, 2010; Liu et al., 2011). As seen in the other studies, the children in the present study found more frequently occurring words with sparse neighbourhoods to be easier to identify. Liu et al. (2011) opined that repeated stimulations with frequently occurring words might have strengthened their memory of the words and consolidate the words in the lexicon compared to the less frequently occurring words. Furthermore, the sparser neighbourhood density might have facilitated retrieval of the 'easy' words as a result of a 'top-down' process. Therefore, children have better mastery of the 'easy' words than the 'hard' words as observed by Liu et al. (2011).

The finding of the present study indicates that the developed material does represent different lexical difficulties. Hence, it can be used to tap the perceptual differences in children and can be used as a valid clinical tool in examining perceptual processes underlying spoken word recognition in Kannada language.

Performance of Children with Hearing Impairment using Cochlear Implants (group II)

The mean and the SD values for the easy and hard words for the 6 participants with hearing impairment using cochlear implants are given in Table 5 for the word scores and Table 6 for the phoneme scores. Similar to the performance of normal hearing group, the children with hearing impairment also performed better on the easy words compared to the hard words. This was observed for both the lists.

Further, Wilcoxon Signed Rank test was done for both word scores and phoneme scores to see if the difference between the easy and the hard words were statistically significant. It was found that there was a significant difference between easy and hard words for list 1 ($Z = 2.232$, $p < 0.05$) and for list 2 ($Z = 2.236$, $p < 0.05$) for word scores. Similarly, for phoneme scores it was found that there was a significant difference between easy and hard words for list 1 ($Z = 2.201$, $p < 0.05$) and for list 2 ($Z = 2.201$, $p < 0.05$).

In addition, Wilcoxon Signed Rank test was also done for both word scores and phoneme scores to see whether there was any difference between the scores obtained across the two lists. The results showed that there was no significant difference between the two lists for the easy words ($Z = 0.04$, $p > 0.05$), hard words ($Z = 0.447$, $p > 0.05$) and the total scores ($Z = 0.378$, $p > 0.05$) for word scores. Likewise, with the phoneme scores there was no significant difference between the two lists for the easy words ($Z = 0.526$, $p > 0.05$), hard words ($Z = 0.949$, $p > 0.05$) and the total scores ($Z = 0.524$, $p > 0.05$). Similar performance across the lists suggests that the two lists developed are equivalent even when used with children using cochlear implant.

This finding of difference in the performance of easy and hard words in cochlear implant users is in consonance with previous studies (Kirk et al., 1995; Kirk, Eisenberg, Martinez & McCutcheon, 1998; Dirks et al., 2001; Yang & Wu, 2005; Yuen et al., 2008; Wang, Wu & Kirk, 2010 & Liu et al., 2011). It suggests that cochlear implant users also utilised their lexical knowledge in word recognition tasks. Kirk et al. (1995) reported that despite a hearing loss and with the degraded sensory input provided via the cochlear implant, their subjects were sensitive to the acoustic-phonetic similarity among the test words. Kirk also reported that though these children have limited vocabularies, they appear to organize words into similarity neighbourhoods in long-term memory, and use this structural information in recognizing isolated words. In a similar manner, the children using cochlear implants in the present study were also able to utilise strategies in a comparable way as that of normal hearing children. This could have lead to them perceive easy and hard words differently.

Comparison of Performance of Children with Normal Hearing and Children with Hearing Impairment using Cochlear Implants

The mean and SD for the word scores and phoneme scores are provided in Figure 1 and Figure 2 respectively. The figures present information of the normal hearing individuals as well as of those with hearing impairment. For both the groups, the mean scores for children with normal hearing were greater than for the children with hearing impairment using cochlear implants.

Mann Whitney U test was used to test the significance of difference between group I and group II for word scores. The results showed that for both the lists there was significant difference between the two groups. This was observed for both easy ($Z = 4.006$, $p < 0.001$) and hard words ($Z= 3.885$, $p < 0.001$) in list 1. Similar findings were obtained in list 2 for the easy ($Z = 4.225$, $p < 0.001$) and hard words ($Z = 3.854$, $p < 0.001$).

To compare the performance of 2 groups on easy and hard words for phoneme scores, Mann Whitney U test was administered which also revealed significant difference between group I and group II. This was observed for both easy ($Z = 3.947$, $p < 0.001$) and hard words ($Z = 3.858$, $p < 0.001$) for list 1. Similar findings were obtained for list 2 for the easy ($Z = 4.163$, $p < 0.001$) and hard words ($Z = 3.884$, $p < 0.001$).

The above finding is in consonance with previous research by Wang, Wu, and Kirk (2010) who studied cochlear implant users and by Patro and Yathiraj (2010) who studied hearing aid users. In both the studies, children with hearing impairment performed significantly poorer than that of normal hearing children on both easy and hard words. Patro and Yathiraj (2010) reported that this poor performance may be because the listening device worn by them was not able to compensate totally for their hearing deficit.

Conclusions

The results indicated a significant effect of lexical difficulty in both the groups (children with normal hearing and for children with hearing impairment using cochlear implants). Both groups performed better on the lexical easy words and poorer on the lexical hard words. There was no significant difference in the performance of children with normal hearing between the two age groups considered. Also, there was no significant difference between the two lists developed thus confirming the inter-list equivalency. The performance of the normal hearing participants was significantly better than that of the cochlear implant users. Similar trend of findings were seen in the performance for both the word scores and phoneme scores for both the groups and for all the parameters.

The material developed in the present study can be used as a valid clinical tool for assessing the speech perception abilities in children. The test can be administered on those children who perform well on simple closed set tests but perform poorly on open set PB word tests, as the difficulty of the test lies in between these two extremes of perceptual difficulty. Also, the test developed can be helpful in selection of appropriate listening devices and may provide guidelines in planning the therapy effectively and also to monitor the progress over time.

References

Alusi, H. A., Hinchcliffe, R., Ingham, B., Knight, J. J., & North, C. (1974). Arabic speech audiometry. *Audiology, 13*(3), 212-220.

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

Begum, A. (2000). *A speech perception test for English speaking hearing impaired Indian preschoolers.* Unpublished Independent project, University of Mysore, Mysore.

Bench, J., Kowal, A., & Bamford, J. M. (1979). The BKB (Bamford Kowal Bench) sentence list for partially hearing children. *British Journal of Audiology, 13*, 108-112.

Berger, K. N., Keating, L. W., & Rose, D. E. (1971). An evaluation of the Kent State University speech discrimination test on subjects with sensorineural loss. *Journal of Auditory Research, 11,* 140-143.

Chowdary, B. K. (2003). *Picture speech Identification test for Hindi speaking children.* Unpublished Master's Dissertation, University of Mysore, Mysore.

Cullington, H. E. (2000). Book review of: Cochlear implant rehabilitation in children and adults by Diane J. Allum. *British Society of Audiology News, 21,* 13.

Dirks, D. D., Takayana, S., & Moshfegh, A. (2001). Effects of lexical factors on word recognition among normal-hearing and hearing-impaired listeners. *Journal of American Academy of Audiology, 12*(5), 233-44.

Elliot, L. L., and Katz, D. (1980). Development of a new children's test of speech discrimination. In F.Martin (Eds.), Hearing disorders in children (pp. 265). Austin, Texas: Prof. Ed. Inc.

Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D. M., & Gantz, B. J. (1992). Performance over time of congenitally deaf and postlingually deafened children using a multichannel cochlear implant. *Journal of Speech and Hearing Research, 35,* 892-902.

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

Haskins, H. (1949). A phonetically balanced test of speech discrimination for children. Unpublished master's thesis, Northwestern University, Evanston, IL.

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 Disorder, 17,* 321-337.

Jerger, J., & Hayes, D. (1971). Audiological manifestations of lesions in the auditory nervous system. *Laryngoscope, 70,* 417-425.

Jerger, J., & Jerger, S. (1974). Auditory findings in brainstem disorders. *Archives of Otolaryngology, 99,* 324-354.

Jijo, P.M., & Yathiraj. A. (2008). *Early speech perception test in Malayalam children.* Published Master's Dissertation, University of Mysore, Mysore.

Kasden, S. D., & Robinson, M. (1970). Otologic-audiologic Hearing Aid Evaluation. *Archives of Otolaryngology, 93*(1), 34-36.

Kirk, K. I., Eiesenberg, L. S., Martinez, A. S. & McCutcheon, M. H. (1998). *The lexical neighbourhood test: Test-retest reliability and interlist equivalency.* Research on spoken language processing, progress report no. 22. Bloomington, IN: Speech Research Laboratory, Indiana University.

Kirk, K. I., Pisoni, D. B., & Osberger M. J. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear and Hearing, 16,* 470-481.

Krull. V., Choi, S., Kirk, K. I., Prusick L., & French, B. (2010). Lexical effects on spoken word recognition in children with normal hearing. *Ear and Hearing, 31*(1), 102-114.

Liu, C., Liu, S., Zhang, N., Yang, Y., Kong, Y., & Zhang, L. (2011). *International Journal of Pediatric Otorhinolaryngology, 75,* 774-781.

Markides, A. (1977). *Binaural hearing aids.* London: Academic Press Inc.

Mayadevi, (1974). *Development and standardization of a common speech discrimination test for Indians.* Unpublished Master's Dissertition, University of Mysore, Mysore.

Miyamoto, R. T., Osberger, M. J., Robbins, A. M., Myres, W. A., Kessler, K., & Pope, M. L. (1992). Longitudinal evaluation of communication skills of children with single- or multichannel cochlear implants. *American Journal of Otology, 13,* 215-222.

Miyamoto, R. T., Osberger, M. J., Todd, S. L., Robbins, A. M., Stroer, B. S., Zimmerman-Phillips, S., & Carney, A. E. (1994). Variables affecting implant performance in children. *Laryngoscope, 104,* 1120-1124.

Moog, J. S., & Geers, A. E. (1990). *Early speech perception test for profoundly hearing impaired* *children.* St. Louis: Central institute for the deaf.

Mukari, S. Z., Ling, L. N., & Ghani, H. A. (2007). Educational performance of pediatric cochlear implant recipients in mainstream classes. *International Journal of pediatrics Otorhinolaryngology, 71,* 231-240.

Patro, C., & Yathiraj, A. (2010). *Lexical Neighbourhood Test: An Indian English version for children.* Published Master's Dissertation, University of Mysore, Mysore.

Prakash, B. (1999). *A picture speech identification test for children in Tamil.* Unpublished Master's Dissertition, University of Mysore, Mysore.

Ritchie, B., and Merklein, R. (1972). An evaluation of the efficiency of the verbal auditory screening for children (VASC). *Journal of Speech and Hearing Research, 15,* 280-286.

Robbins, A. M., Renshaw, J. J., & Osberger, M. J. (1995). *Common phrase test.* Indianapolis, IN: Indiana University School of Medicine.

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

Rout, A. (1996). *Perception of monosyllabic words in Indian children*. Unpublished Master's Dissertation, University of Mysore, Mysore.

Singh, S. (1966). Cross language study of perceptual confusion of plosives in two conditions of distortion. *Journal of the Acoustical Society of America, 40,* 635-656.

Singh, D. K. (2010). *Development of multi-syllabic lexical neighbourhood test in Hindi.* Unpublished Master's dissertation, University of Bangalore, Bangalore.

Staller, S. J., Beiter, A. L., Brimacombe, J. A., Mecklenburg, D. J., & Arndt, P. (1991). Pediatric performance with the Nucleus 22-channel cochlear implant system. *American Journal of Otology, 12,* 126-136.

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

UNICEF Research project by RRTC and AYJNIHH. (1990). *Development and standardization of language and articulation tests in seven different languages.*

Vandana, S. (1998). *Speech identification test For Kannada speaking children.* Unpublished Independent project, University of Mysore, Mysore.

Wang, N. M., Wu, C. M., & Kirk, K. I. (2010). Lexical effects on spoken word recognition performance among Mandarin speaking children with normal hearing and cochlear implants. *International Journal of Pediatric Otorhinolaryngology, 74,* 883-890.

Yang, H. M., & Wu, J. L. (2005). Mandarin Lexical Neighborhood Test (M-LNT) for pre-school children: development of test and its validation, *Journal of Taiwan Otolaryngology-Head and Neck Surgery, 40*(1), 1-12.

Yuen, K. C., Ng, I. H., Luk, B. P., Chan, S. K., Chan, S. C., Kwok, I. C., Yu, H.C. et al. (2008). The development of Cantonese Lexical Neighborhood Test - a pilot study. *International Journal of Pediatric Otorhinolaryngology, 72,* 1121-1129.

Effects of Compression Release Time in Hearing Aid on Acoustic and Behavioural Measures of Speech

[1]Arpitha V. & [2]Manjula P.

Abstract

The present study investigated the effects of varying the release time (40 ms, 640 ms & 1280 ms) on acoustic and behavioural measures of speech. The acoustic measures included a temporal measure or Envelope Difference Index (EDI) and a spectral measure or Log Likelihood Ratio (LLR). These were computed for VCV tokens of six classes of speech sounds. The behavioural measures were Speech Identification scores (SIS) and quality ratings for loudness, clarity, naturalness and overall impression. The effect of the release time was investigated at three input levels (30, 45 & 65 dBHL). The behavioural measures were performed on Group A participants (N=10; with moderate to moderately severe HL) and Group B participants (N=10; with moderately severe to severe HL). The acoustic measures were performed on five individuals (hearing loss above moderately severe) using the mean audiometric data of Group A and Group B. The effect of release time on EDI and LLR varied at each input level. At low input level of 30 dBHL, the effect of release time was not significant on all speech sounds. At 45 and 65 dBHL, shorter release time (40 ms) showed significantly higher temporal and spectral distortions than longer release times (640 and 1280 ms) for majority of speech sounds. The amount of temporal and spectral distortions was found to be proportional to input level. SIS and quality ratings increased as the release time was made longer. Although not significant, Group B showed higher distortions and lesser SIS than Group A.

Keywords:*Release time, Input level, Envelope difference index (EDI), Log likelihood index (LLR), SIS and quality rating*

Introduction

A cochlear hearing loss implies a poorer functioning of the compressive non-linearity mechanism. Auditory dysfunction due to a loss in cochlear non-linearity can be described in several psycho-acoustical terms, such as a reduced spectral and temporal resolution, a disturbed loudness perception and an increased temporal and spectral masking (Oxenham & Bacon, 2003). Hence, to combat such problems, dynamic range compression has been introduced in hearing aids. As the level of the signal is enhanced, the spectral resolution becomes more impaired and this also leads to spread of masking. Hence, it can be inferred that individuals with hearing impairment rely more on temporal cues than spectral cues. Thus, it is necessary to maintain the temporal cues of the speech processed by hearing aid in order to compensate for their loss. In the WDRC compression system, the major factors influencing the temporal information of speech envelope include compression time constants i.e., attack and release time (Souza, 2002).

With respect to compression time constants, if the release time is too short, then the gain will vary during each voice pitch period, i.e., the compressor will distort the waveform. If the release time is made longer, rapid gain fluctuations will be reduced and thus the distortions would be minimal. However, for the intense signal which is of brief duration, short attack time causes rapid reduction in gain. But, long release time has its undesirable effect by having the gain to remain low for a longer time after a brief sound (Kuk, 1996). During compression activation or de-activation, due to the brief attack or release time intervals, the spread of energy and the distortions induced are inevitable, contributing to the temporal and spectral artifacts (Wang, 2001)

Research so far, however, is inconclusive in providing guidelines on the correct setting of release time. Literature reviews on compression provide no compelling answers with regard to release time (Dillon, 1996; Hickson, 1994). This is due to a number of variables compounding the findings in such studies. Two approaches have been used to study the release time in compression hearing aids. They are studying the effects of this parameter on speech intelligibility, and on quality or user preference. With regard to intelligibility, a few available studies conflict with each other showing either no difference among settings of different release times (Bentler & Nelson, 1997) or that moderate values (between 60 and 150 ms) provide better intelligibility than very short or very long release times (Walker & Dillon, 1982). Studies of user preference showed mixed results; some show no user preferences for any release time (Bentler & Nelson, 1997), while others rated longer release times to be more pleasant than shorter release times (Hansen, 2002)

To address the issue of an optimal release time for hearing aid fittings, it is necessary for a comprehensive analysis of the acoustic effects of compression on speech and the resulting effects on speech perception. The acoustic analysis must focus on the aspects of speech

[1]Email:Arpitha_v25@yahoo.com
[2]Professor in Audiology, Email: manjulap21@aiishmysore.in

that are mostly likely to be affected by varying the release time such as, temporal and spectral parameters.

Different acoustic measures are available to quantify the effect of changes in the acoustic properties of the signal. They include measures such as Envelope Difference Index (EDI), CVR (Consonant-to-Vowel Ratio), and Spectral Distance Measure (SDM). The EDI is an index of change in the temporal envelope between two signals. The value ranges from 0 to 1, where 1 represents completely dissimilar envelopes and 0 represents perfectly similar envelopes (Fortune, Woodruff, & Preves, 1994). This technique is based on envelope subtraction, and generates an Envelope Difference Index that may help to detect the alteration of the natural speech envelope due to amplification. The Log-Likelihood Ratio (LLR) is a spectral distance measure which mainly models the mismatch between the formants of the original and enhanced signals (Rao, Murthy, & Rao, 2011). Jeon and Lee (2008) have used LLR for objectively analyzing the spectral distortions caused due to spectral enhancement technique in hearing aids and found that the LLR values increased as the weightage of the enhancement increased.

Amongst the approaches used, there is a dearth of studies on acoustic effects of compression release time. Also a few studies on behavioural parameters have used different release times with a single input level (Neuman, Bakke, Mackersie, Hellman, & Levitt, 1995; Novick, Bentler, Dittberner, & Flamme, 2001). There is also a need for different settings in compression parameters which maintains the temporal cues for speech recognition, for individuals with different degrees of loss (Van Tasell, 1993).

The aim of the present study was to investigate the effect of compression release time (40 ms, 640 ms and 1280 ms) on temporal and spectral characteristics of the VCV tokens in the aided condition, at three different presentation levels (30, 45 and 65 dB HL). Also, to evaluate the effect of different compression release times (40 ms, 640 ms & 1280 ms) on the speech intelligibility and perceived quality judgments of speech through the hearing aid set at three different presentation levels (30, 45 and 65 dB HL), in individuals with hearing impairment.

Method

The aim of the study was to evaluate the acoustic and behavioural effects of different release times of compression in a hearing aid. The following procedure was used.

Participants for Behavioural Measures

Thirteen male and seven female participants were considered in the study. Their age range was from 20 years to 55 years (mean: 42.9 years, SD8.76). All the 20 participants were diagnosed to have flat or gradually sloping sensorineural hearing loss (SNHL) ranging from moderate to severe degree. The participants were divided into two groups based on the degree of hearing loss. Group A consisted of ten participants having moderate to moderately severe SNHL. Their Pure Tone Average (PTA) ranged from 41 dB HL to 70 dB HL (mean: 60.9, SD: 7.89). Group B consisted of ten participants having moderately severe to severe SNHL. Their PTA ranged from 71 dB HL to 90 dB HL (mean: 75.4, SD: 4.57). Their speech recognition thresholds and Speech Identification Scores were proportionate to their hearing thresholds. All the participants had normal middle ear status. The participants had post-lingually acquired hearing loss with age adequate speech and language. They did not have any previous history of neurological and cognitive dysfunction. Out of twenty participants, twelve participants had more than three months of experience with the hearing aid usage; the other eight participants were naive hearing aid users.

Participants for Acoustic Measurement

Five participants having moderately-severe to severe hearing loss were considered. Their age ranged from 20 to 30 years (Mean: 24.2 years, SD: 2.48). Acoustic measures were performed on these five participants with the hearing aid programmed for the mean audiometric data of Group A (mean PTA 60.9) and Group B (mean PTA of 75.3 dB HL).

Preparation of VCV Tokens and Kannada Story Passage

Three female speakers whose mother tongue was Kannada were chosen to utter the Vowel-Consonant-Vowel (VCV) tokens. The phonemes in Kannada with greater than 0.5% of frequency of occurrence (Ramkrishna, Nair, Chiplunkar, Atal, Ramachandran, & Subramanian, 1962) were selected and paired with a combination of vowel /i/. Table 1 gives the seven speech sounds used in the study, classified based on manner of articulation. The VCV tokens uttered by the three speakers were recorded on to a computer, using the Adobe Audition software, via the recording microphone (Ahuja, AUD-101XLR) placed at a distance of 10 cm from the lips of the speaker (Winholtz & Titze, 1997). The recorded stimulus was digitized using a 32-bit processor at 44,100 Hz sampling frequency. The recorded 21 syllables were subjected to goodness test in order to select one set of test stimuli. These test stimuli were then concatenated with an inter-stimulus interval of 3 sec. These stimuli were preceded by a 1 kHz calibration tone and were written on to a compact disc. Similarly, one of the speakers read out the Kannada passage and this was recorded using the above mentioned procedure. The stimulus was recorded in an air conditioned sound treated room.

Table 1: Classification of consonants based on manner of articulation

Stops	/ipi/, /iti/, /iki/, /ibi/, /idi/, /igi/, /iti/, /idi/
Nasals	/ini/, /imi/, /ini/
Fricatives	/isi/, /ishi/, /ihi/
Affricates	/ichi/, /idzi/
Liquids	/iri/, /ili/, /illi/
Glides	/ivi/, /iyi/

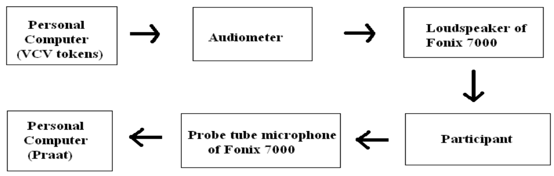


Figure 1: Block diagram depicting the set-up of equipment for acoustic data collection.

Hearing Aid Programming

For the purpose of the study, a digital behind the ear hearing aid with an option for varying its compression release time, attack time, phase cancellation feedback reduction technique, two listening programmes, two channels and seven bands was used. This hearing aid was programmed using the NAL-NL1 prescriptive formula, with acclimatization level of 2 via NOAH software and was optimized for each individual. WDRC with compression threshold of 50 dB SPL was selected. Other than release time, all other parameters were kept constant across conditions. Electroacoustical measurements were performed to verify the parameters set in the hearing aid.

Acoustic and Behavioural Measurements

Acoustic (unaided and aided) and behavioural (aided) measurements were carried out with the release time set to 40 ms, 640 ms and 1280 ms at three different input levels. The data were collected in two phases; Phase I, Measurement of acoustic parameters - temporal (EDI) and spectral (LLR) parameters and Phase II, Measurement of behavioural parameters - speech identification scores (SIS) and speech quality ratings.

Phase I: Measurement of acoustic parameters - temporal and spectral

The acoustic measures used in the study were Envelope Difference Index (EDI) and a Spectral Distance / Distortion Measure (SDM) called Log-Likelihood Ratio (LLR). The following steps were involved in measuring these two parameters.

Step A: Set-up for acoustic measurement: A personal computer containing the VCV tokens was connected to the auxiliary input of the portable diagnostic audiometer. The output of the audiometer was routed to the loudspeaker of the real ear measurement system, Fonix 7000. The input from the Fonix 7000 was disabled and the input from the audiometer was routed to the loud speaker. The probe tube microphone was inserted into the unoccluded ear canal of the participant. The output from the Fonix 7000 system was routed to the microphone inlet of the PC containing Praat software. The participants were seated comfortably on a chair at a distance of 12 inches away and at 45^0 Azimuth from the loudspeaker of the real ear measurement system. Figure 1 shows the set-up of the equipment used for recording.

Step B: Unaided measurement for acoustic analysis: The probe tube microphone of the real ear measurement system was inserted in the ear canal. The length of the probe tube inserted in the ear canal was 3-5 mm beyond the length of the ear mould of the participant. The VCV tokens were presented to the participant. This was picked up from the ear canal of the participant and recorded on to the PC using Praat software, for further analysis. This procedure was repeated at three presentation levels, i.e., 30, 45 and 65 dB HL.

Step C: Aided measurement for acoustic analysis: Hearing aid was set to one of the three release times and was fitted using custom ear mould. The probe tube microphone was inserted such that the tube was 3 to 5 mm beyond the ear mould. The VCV tokens were presented from the loudspeaker and were picked up by the hearing aid; the output of the hearing aid in the ear canal was in turn picked up by the probe tube microphone of the real ear measurement system. The signal from the probe tube microphone was recorded in Praat software in the computer for further analysis. This procedure was performed for three release times (40 ms, 640 ms & 1280 ms) and at three input levels (30, 45 & 65 dB HL). The aided probe tube microphone measurements were done on five participants for the mean audiological data of two groups of participants (Group A and Group B of behavioural measurement).

Step D: Calculation of EDI and LLR: Using the MATLAB functions, different algorithms or codes were generated to calculate the EDI and LLR values. To calculate EDI and LLR, the speech VCV tokens (unaided and aided) which have to be compared were edited with respect to the common reference point shared by them. The edited aided and unaided VCV tokens were then fed into the MATLAB software containing particular algorithm for the calculation of EDI and LLR values.

Phase II: Measurement of behavioural parameters - Speech Identification Scores and speech quality rating

Behavioural measures used in the study were speech identification scores (SIS) and speech quality judgment. For both the measurements in aided condition, the hearing aid was programmed as mentioned earlier.

Step A: Measurement of aided SIS: The participants were seated comfortably on a chair, one meter away from the loudspeaker of the audiometer, at 45^0 Azimuth. They were tested for speech identification using bi-syllabic words (Yathiraj & Vijayalakshmi, 2005). This word-list included five lists and each list comprising of 20 words, which are phonemically balanced. The words were presented in monitored live mode through the loud speaker of the sound field diagnostic audiometer. This was carried out for different release time (40 ms, 640 ms & 1280 ms) at three input levels (30, 45 & 65 dB HL) on two groups of participants. The participants were instructed to repeat the words. In each condition, the number of words correctly identified was noted.

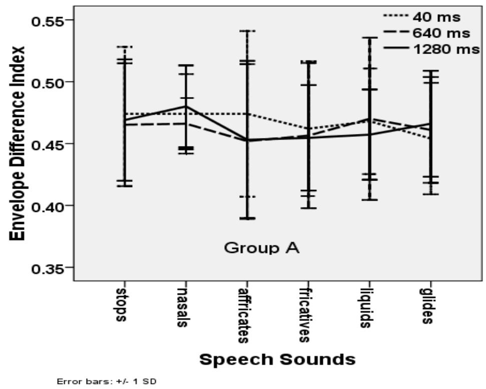


Figure 2

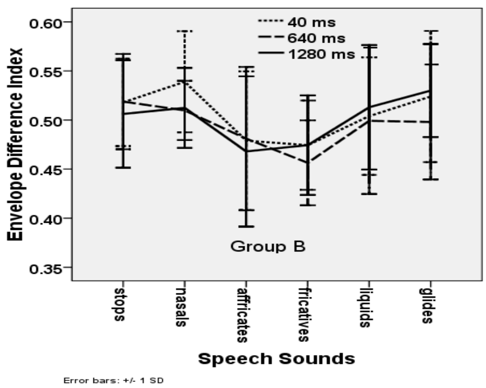


Figure 3

Figures 2 and 3: Effect of three release times at 30 dB HL input level on EDI of six classes of speech sounds for mean audiometric data of Group A and Group B, respectively.

Step B: Measurement of aided speech quality judgment: To assess the speech quality, a recorded Kannada story passage was presented via recorded mode through the loud speaker of the sound field diagnostic audiometer. The participants were instructed to rate the story for its quality on a 10-point rating scale for the parameters of loudness, clarity, naturalness and overall impression. The rating scale ranges from 0 to 10 wherein, 0 indicates 'very poor' 2 indicates 'poor' 4 indicates 'fair' 6 indicates 'good', 8 indicates 'very good' and 10 indicates 'excellent'. The participants were instructed to use odd numbers for rating when the perception was between any the two points. This rating procedure was carried out for different release time (40 ms, 640 ms & 1280 ms) at three input levels (30, 45 & 65 dB HL) on two groups of participants.

For the computation of EDI, the method used by Fortune, Woodruff, and Preves (1994) was used. For computation of LLR, COLEA, a software tool for speech analysis developed by Loizou (1998) was used. The procedure adopted to compute LLR was the same as the one used by Jeon and Lee (2008); Cote, Turbin, and Moller (2008).

Results and Discussion

Appropriate statistical analysis was carried out to analyze the data. The results of the study are discussed for acoustic (EDI & LLR) and behavioural measures (SIS & quality rating) of speech.

Acoustic Measures

Non-parametric statistics was used for the analysis as the measurement was done on five participants, to evaluate the effect of release time on EDI and LLR.

Envelope Difference Index (EDI)

Descriptive statistics was performed to compute the mean and standard deviation (SD) of EDI values for six classes of speech sounds viz., stops, nasals, affricates, fricatives, liquids, and glides, at three release times and at three input levels (Figure2, 3, 4, 5, 6 & 7) on the data of two groups. A test of two independent samples i.e., Man Whitney U test, was performed in order to compare between groups. Two related samples test, i.e., Wilcoxon Signed rank test was performed in order to compare across conditions (Table 2).

Effect of release times on six groups of speech sounds: The results of comparison across three release time conditions are discussed below separately at each input level.

At 30 dBHL: As depicted in Figures 2 and 3, there is no systematic trend seen for the mean EDI values, as the release time increased. As in Table 3, Wilcoxon Signed Rank test revealed no significant difference between release times ($p>0.05$). This trend was noticed for all the six groups of speech sounds and for both the groups.

The results are in consonance with that reported by Jenstad and Souza (2005). According to their study, at lower input levels of 50 dB SPL, the EDI values underwent minimal changes across release times used (12 ms,

Table 2: Significance of effect of release time on EDI for six classes of speech sounds at three input levels for group A and Group B mean audiometric data

Pair wise comparison of EDI for Group A

Input level	***30 dB HL***			***45 dB HL***			***65 dB HL***		
Release time	40 & 640	40 & 1280	640 & 1280	40 & 640	40 & 1280	640 & 1280	40 & 640	40 & 1280	640 & 1280
Stops					*		*	*	
Nasals				*	*		*	*	*
Affricates							*	*	*
Fricatives								*	
Liquids								*	
Glides				*	*	*		*	*

Pair wise comparison of EDI for Group B

Input level	***30 dB HL***			***45 dB HL***			***65 dB HL***		
Stops							*	*	
Nasals					*		*	*	*
Affricates				*	*	*	*	*	*
Fricatives							*	*	*
Liquids							*	*	
Glides							*	*	*

*Note: *significant difference at p<0.05 At 30 dB HL*

100 ms & 800 ms). This can be explained by the fact that at lowest input level of 30 dB HL used in the study, the signal would be below the compression threshold. If the signal falls below the compression threshold, that particular signal would be amplified linearly rather than non-linearly. Hence, the effect of release times at this low input level is not significant and also shows larger variability.

At 45 dB HL: As depicted in Figure 4 and Figure 5, the mean EDI values decreased as the release time increased for all six classes of speech sounds. The significant difference between release times (Table 2) are not constant across speech sound classes. On observation, the EDI values at 45 dB HL increased in comparison to 30 dB HL for all the conditions. The similar trend is seen for both groups.

The results of the present study are in agreement with that carried out by Jenstad and Souza (2005). Longer release time had significantly reduced temporal envelope distortions compared to short release time. This can be due to the fact explained by Kuk (1996) that at short release time (40 ms used in the current study), the gain will vary during each voice pitch period, and hence the compressor will distort the waveform. If the release time is made longer, rapid gain fluctuations will be reduced and thus the distortions would be minimal.

The results also revealed variable effects of release time on different consonants. A few consonants had significantly reduced EDI and other consonants did not vary significantly. Although not significant for few of the speech sounds, the mean EDI values for all sounds decreased as the release time increased from 40 ms to 640 ms and to 1280 ms.

The mean EDI values in most of the test conditions revealed that the temporal envelope distortion of consonants is higher for glides, stops, affricates, nasals, liquids and fricatives, in that order. This is because, the greatest effect is on sounds where critical information is carried by variations in sound amplitude over time, such as stops which can be modeled as a series of temporal cues with falling or raising burst spectrum, onset of voicing, and air release. Speech sounds such as stops, affricates and glides are known to have faster temporal variations, i.e., sharp rise time in burst of stops and affricates, faster transition in glides in contrast to semivowels, as compared to other classes of speech sounds such as fricatives which are high frequency hiss and are longer in duration. They vary slowly in terms of temporal parameters (Savithri, 1989).

At 65 dB HL: As depicted in Figure 6 and Figure 7, the mean EDI values decreased as the release time increased for all the six classes of speech sounds. The significant difference between release times (Table 2) are not constant across speech sound classes.

At 65 dBHL, longer release time of 1280 ms produced significantly reduced distortions for majority of speech sounds than at 45 dBHL. On observation, the EDI values at 65 dB HL increased in comparison to 45 dB HL for all the conditions. The similar trend is seen for both the groups.

As depicted in Table 2, at 65 dBHL, the effects of release time on temporal envelopes are significant for most of the speech sounds compared to 45 dBHL. This implies that, at higher input levels, the compression is more effective (Henning & Bentler, 2008), i.e., the more the input intensity, more is the reduction in gain provided. As described by Neuman et al. (1996), longer release time play a major role to offset the effects of increased compression, as the compression increases.

Hence, it can be inferred from the findings that the sig-

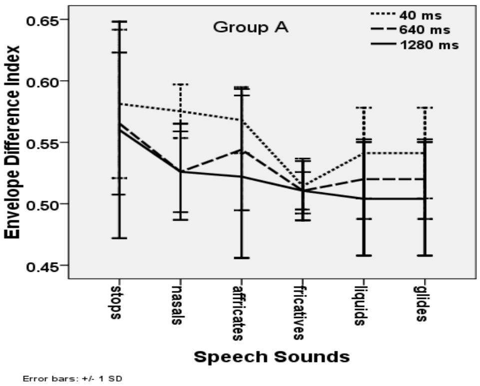


Figure 4

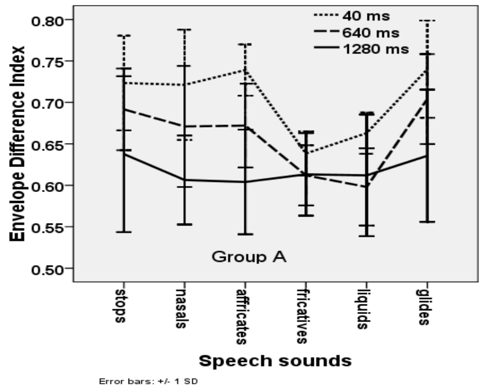


Figure 6

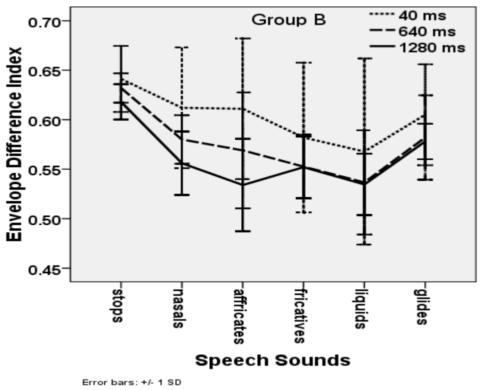


Figure 5

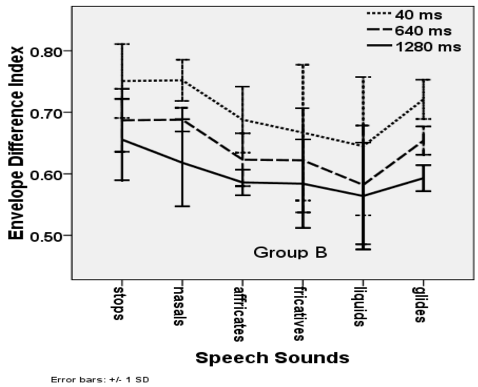


Figure 7

Figures 4 and 5: Effects of three release times at 45 dB HL input level on EDI of six classes of speech sounds for mean audiometric data of Group A and Group B.

Figures 6 and 7: Effect of three release times at 65 dB HL input level on EDI of six classes of speech sounds for mean audiometric data of Group A and Group B.

nificant effect of short and long release times on temporal envelope distortion is more evident at higher input levels. As discussed earlier, on general observation of the mean EDI values, the temporal envelope distortion of consonants are higher for glides, stops, nasals, affricates, liquids and fricatives, in that order. This is due to the same fact as explained earlier for effect of release time at 45 dB HL.

Effect of degree of hearing loss on EDI: On observation of the mean EDI values, the difference between the two groups across the conditions may be attributed to the fact that, the dynamic range into which the signals have to be compressed was less for Group B than for Group A. Hence, the compression ratio to fit the entire intensity range within the dynamic range was set to be higher. At higher compression ratio almost all sounds will undergo effective compression and the role of shorter and longer release times are more evident under this condition (Henning & Bentler, 2008; Neuman et al 1996). The two groups are not statistically different as the mean audiometric thresholds of Group A and Group B differs only by 10 to 15 dB HL.

Log-Likelihood Ratio

Descriptive statistics was performed to compute the mean and standard deviation (SD) of LLR values for six classes of speech sounds viz., stops, nasals, affricates, fricatives, liquids, and glides, at three release times and at three input levels (Figure, 8, 9, 10, 11, 12, 13) on the data of two groups. Man Whitney U test, a test of two independent samples, was performed in order to compare between groups. In order to compare across conditions two related samples test i.e., Wilcoxon Signed rank test, was performed (Table 3).

Effect of release time on Log-Likelihood Ratio: The results of comparison across three release time conditions are discussed below, at each input level separately.

The results of LLR on six classes of speech sounds at each of the level across three release times follow the same trend as seen in the effects of release times on EDI. The mean LLR values at 30 dBHL do not vary systematically with the release time. As depicted in the Table 3, and Figures 8 and 9, at 30 dBHL, there was no significant effect of release time. As depicted in the respective figures, the mean LLR values decreased as the release time increased, at 45 dBHL and 65 dBHL. At 45 dBHL and 65 dBHL, the spectral distortion was significantly higher at 40 ms release time condition and as the release time increases from 640 to 1280, the spectral distortions tend to decrease (Table 3, Figure 10, Figure

Table 3: Significance of effect of release time on LLR for six classes of speech sounds at three input levels for group A and Group B mean audiometric data

Pair wise comparison of EDI for Group A									
Input level	***30 dB HL***			***45 dB HL***			***65 dB HL***		
Release time	40 & 640	40 & 1280	640 & 1280	40 & 640	40 & 1280	640 & 1280	40 & 640	40 & 1280	640 & 1280
Stops		*	*						*
Nasals			*					*	*
Affricates					*		*	*	
Fricatives							*	*	*
Liquids					*	*	*	*	
Glides					*	*	*	*	*
Pair wise comparison of EDI for Group B									
	30 dB HL			***45 dB HL***			***65 dB HL***		
Stops		*			*	*			
Nasals			*	*	*		*	*	*
Affricates									
Fricatives		*		*			*	*	*
Liquids							*	*	
Glides				*	*	*	*	*	

*Note: *significant difference at p<0.05 At 30 dB HL*

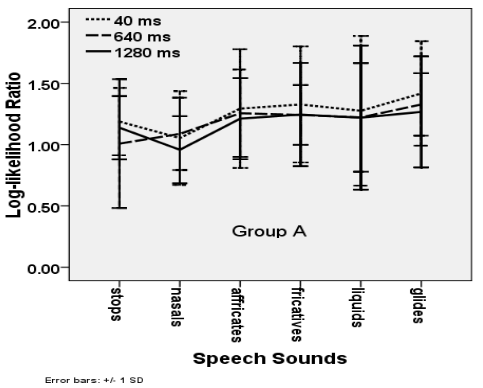


Figure 8

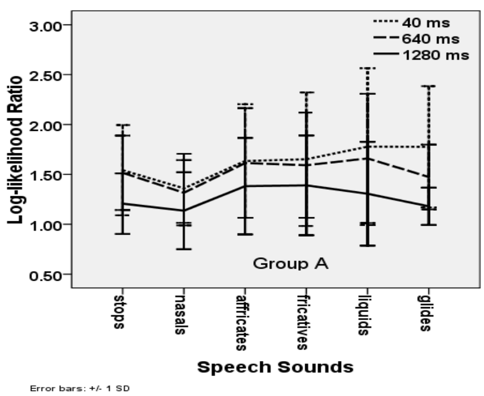


Figure 10

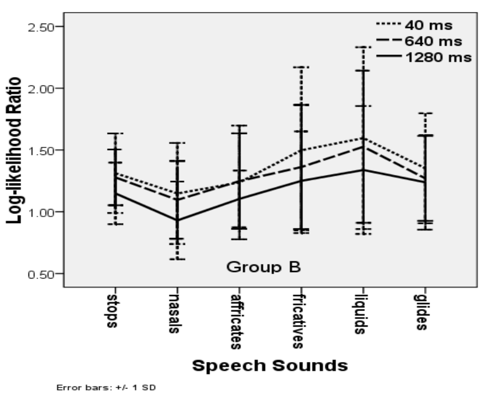


Figure 9

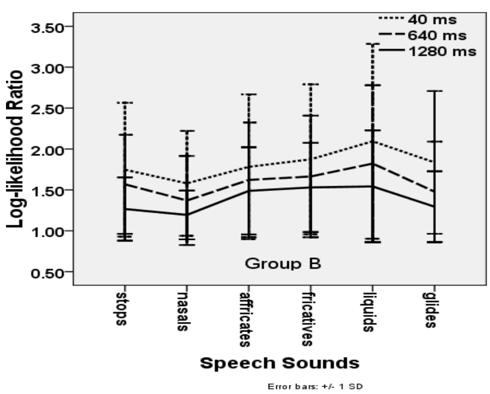


Figure 11

Figures 8 and 9: Effect of three release times at 30 dB HL input level on LLR of six classes of speech sounds for mean audiometric data of Group A and Group B.

Figures 10 and 11: Effect of three release times at 45 dB HL input level on LLR of six classes of speech sounds for mean audiometric data of Group A and Group B.

11, Figure 12 & Figure 13). This is true for both 45 and 65 dBHL input levels. But, as depicted in the Table 3, the effect of release time on LLR was significantly greater at 65 dB HL than at 45 dB HL condition. Also, the majority of speech sounds shows significant effect of release times at 65 dB HL, compared to 45 dB HL input level. As it was seen for EDI, irrespective of the input level and RT, not all speech sounds are affected to the same extent. There exists a lot of variability.

In general, the overall results of LLR are similar to that of EDI, i.e., there are higher spectral distortions at short

(40 ms) release time. In addition, a spectral distortion significantly reduces as the release time increases from 640 ms to 1280 ms. This can be explained based on the fact that, alterations in one of the domains (temporal/spectral) nearly always have corresponding effects in the other one. That is, changes to the temporal waveform of the envelope produce corresponding spectral changes, and vice versa (Van Tassel, 1993).

The result of shorter release time having greater spectral distortions can be attributed to the reason described by Wang (2001). When the release times are made shorter, they are associated with the broadest and strongest distortion across the frequency spectrum due to the spread of energy. On the other hand, as the release times are made longer, they are associated with the narrowest and weakest distortion on the frequency spectrum of the signal.

On close observation of mean LLR values, it was noted that the spectral distortion of consonants can be placed in order of higher to lower amount of distortions as in liquids & fricatives, affricates, glides, stops and nasals. A similar trend was noted for both the mean audiometric group data. Unlike EDI, which is a time and intensity based measure, the LLR compares formants of speech tokens at every fixed time frame and displays one single value. Hence, both the measures tap different aspects in a particular signal. The speech sounds which are confounded majorly by temporal variations tend to be more vulnerable to changes detected in the EDI than in LLR and vice versa.

The order in which the consonants have more spectral distortion can be due to the fact that, the formant frequencies of consonants range from low to high frequencies. For example, in Kannada language, formant frequencies of velar stops range from 400 Hz to 1.9 kHz, liquids have their formants from 450 Hz to 2.9 kHz, nasals have their formants from 300Hz to 1.5 kHz, formants of fricatives ranges from 500 Hz to 4.9 kHz etc. The compression amplification used in the study is a two channel device and each of the channel vary in terms of compression ratios. Hence, each portion of the speech signal might have undergone varied degrees of compression resulting in diffused effects of spectral distortion among the six classes of speech sounds (Savithri, 1989).

The magnitude of the temporal/spectral artifacts during compression activation/deactivation is not constant, but constantly changing. Specifically, broader and stronger distortion across frequency spectrum occurs within the shorter attack/release times whenever the input signal level rises to or falls from slightly above the compression threshold (Wang, 2001). In simpler words, at higher levels of input in contrast to low levels, maximum compression takes place and hence more spectral distortions are noted.

Table 4: Mean and SD of SIS across three input levels and across three release times for Group A and Group B

Input Level (in dB HL)	Release time (ms)	Group A	Group B
		Mean (SD)	Mean (SD)
30	40 ms	14.70 (0.68)	11.90 (0.88)
	640 ms	15.70 (0.82)	11.60 (1.08)
	1280 ms	15.50 (1.08)	12.50 (0.71)
45	40 ms	17.50 (0.53)	15.00 (0.82)
	640 ms	18.00 (0.67)	15.00 (0.82)
	1280 ms	18.10 (0.74)	15.70 (0.68)
65	40 ms	19.10 (0.32)	16.80 (0.92)
	640 ms	19.10 (1.10)	16.90 (1.40)
	1280 ms	19.10 (1.19)	17.20 (0.92)

Effect of degrees of hearing loss on LLR: On observation of the mean LLR values in Table 4, although not significant, the mean LLR values were higher for Group B audiometric data. The difference between the two groups across the conditions may be attributed to the fact that, group B had lesser dynamic range and hence more compression ratio caused more spectral distortions. Hence at higher compression ratio, almost all sounds will undergo effective compression and the role of shorter and longer release times are more evident under this condition (Henning & Bentler, 2008; Neuman et al. 1996). The fact that the two groups are not statistically significant is because the mean audiometric thresholds of Group A and Group B differ only by 10 to 15 dB HL.

A few of the varied results noted with respect to speech sounds of a particular category across release times may be because of the reason that, within each class, the speech sounds constituted of both voiced and voiceless consonants. These effects of release time for voiced sounds might be different from that of a voiceless speech sound. Hence, it would be important to study the effect of release time on speech sounds categorized based on place of articulation, manner of articulation and based on voicing characteristics. However, due to time constraints, the study was carried out only based on categorization of manner of articulation.

Behavioural Measures

The behavioural measures, SIS and quality ratings, across conditions are discussed for Group A and Group B.

Speech Identification Scores

Descriptive statistics was performed to compute the mean and standard deviation (SD) of SIS. Table 4 depicts the mean and SD of SIS, across three release times and across three input levels, for Group A and Group B. Two way repeated measure ANOVA revealed that the

Table 5: Effect of three release times across three input levels on perceptual ratings of quality for Group B

Release time (ms)	***Input Level (dBHL)***	***Group A - Quality ratings***			
		Loudness	***Clarity***	***Naturalness***	***Overall Impression***
		Mean (S.D)	***Mean (S.D)***	***Mean (S.D)***	***Mean (S.D)***
40	30	5.20(1.31)	5.00(1.33)	4.60(1.21)	5.20(1.13)
	45	6.90(1.28)	6.50(1.26)	6.00(1.41)	6.90(1.37)
	65	8.50(1.08)	7.60(1.50)	6.30(1.41)	8.30(0.91)
640	30	4.90(1.30)	5.50(1.17)	5.20(0.63)	5.50(1.64)
	45	6.60(1.56)	6.90(1.10)	7.20(1.03)	7.40(0.96)
	65	8.15(2.10)	8.20(1.31)	7.40(1.07)	8.60(1.42)
1280	30	5.10(1.10)	5.80(1.22)	5.90(0.99)	6.00(0.94)
	45	6.40(1.26)	7.80(1.39)	7.40(0.96)	8.10(0.96)
	65	8.60(1.76)	8.80(1.31)	8.20(0.91)	9.20(1.00)

interaction between group and conditions are statistically significant [$F(1, 64.14) = 3.283$, $p<0.05$]. In addition, significant main effect of conditions [$F(8, 144) = 119.56$, $p<0.05$] was also present. Bonferroni's multiple comparison test results are discussed for Group A and B.

Effects of Release Time on SIS: Table 4 depicts the mean and SD of SIS across release times and across three input levels for both Group A and Group B.

Group A: As depicted in Table 4, the mean scores are higher for longer release times, at both 30 dB HL and 45 dB HL. At 30 dBHL, repeated measure ANOVA revealed that the release times were significantly different [$F(2, 18)$ 7.132: $p<0.05$]. Pair-wise differences with Bonferroni's multiple comparison showed a significant difference between 40 and 640 ms, at 0.05 level of significance.

At 45 dBHL, repeated measure ANOVA revealed that the release times were significantly different [$F(2, 18)$ 7.154: $p<0.05$]. Bonferroni's multiple comparison results showed a significant difference between 40 and 640 ms and also between 40 and 1280 ms, at 0.05 level of significance.

The high speech intelligibility scores obtained at longer release times and significantly low SIS at short (40 ms) release time is because, short release time causes unnatural alteration of vowel and consonant ratio (Freyman, Nerbonne, & Cote, 1991). At higher levels of 65 dBHL, there was no improvement in SIS as the release time increased. A study by Vanaja and Jayaram (2006) reported the mean SIS for different degrees of hearing. According to their study, listeners with moderate hearing loss would have mean SIS of around 85 % (12.47 SD) and listeners with moderately severe hearing loss would have mean SIS of around 77.5 % (13.89 SD). For listeners in the current study, the audibility was compensated through hearing aid, and additional audibility was provided at the higher input level of 65 dB HL. The SIS at this level in the current study reached its maximum performance, and the effect reached plateau across release time. This could be attributed to the reason of ceiling effect of SIS at 65 dB HL.

Group B: As depicted in the Table 4, the mean SIS slightly increased as the release time increased at each input level. Repeated measures ANOVA revealed no significant difference across RT conditions in any of the input levels. Although not significant, the mean SIS increased for longer release times (1280 ms) at 45 and 65 dB HL. This could be attributed to the fact that Group B listeners had very narrow dynamic range and hence the signal had to be compressed to a larger extent. When there is more compression due to high gain, high compression ratio or higher input level, longer release time can offset the possible temporal distortions to some extent compared to shorter release times (Henning & Bentler, 2008). The finding of difference not being significant can be due to the reason that listeners used other cues like contextual cues, rather than depending upon only temporal ones (Jenstad & Souza, 2005).

Perceptual Quality Ratings

Descriptive statistics was done to compute the mean and standard deviation of quality ratings for four parameters. Repeated measures ANOVA was administered to find out the overall interaction of the conditions and groups. Bonferroni's multiple comparison test was done to determine the significantly different pairs. Multivariate Analysis of Variance (MANOVA) was carried out to compare between the two groups. As depicted in Table 5 and Table 6, the mean and SD for loudness, clarity, naturalness and overall impression parameters, across three release times and across three input levels, for Group A and Group B. The results are described below for each of the parameter, at each input level, for the two groups separately.

Table 6: Effect of three release times across three input levels on perceptual ratings of quality for Group B

Release time	Input Level(dBHL)	Group B - Quality ratings			
		Loudness	Clarity	Naturalness	Overall Impression
		Mean (S.D)	Mean (S.D)	Mean (S.D)	Mean (S.D)
40	30	4.20(0.76)	4.50(1.08)	4.65(0.88)	4.60(0.69)
	45	6.30(1.39)	5.80(1.54)	5.65(1.29)	6.50(1.17)
	65	8.70(0.94)	7.20(1.39)	6.65(1.29)	8.00(0.47)
640	30	4.00(0.91)	4.70(0.63)	5.10(0.56)	5.25(0.97)
	45	6.65(0.88)	6.50(1.08)	6.90(0.73)	7.30(0.91)
	65	8.80(0.78)	7.55(1.77)	7.15(1.41)	8.35(1.20)
280	30	4.00(0.93)	4.85(0.66)	5.45(0.72)	5.75(0.85)
	45	6.30(0.94)	7.20(0.42)	7.25(0.54)	7.45(0.76)
	65	8.40(0.69)	7.65(1.70)	7.60(1.17)	8.40(0.93)

Table 7: Summary of results of perceptual quality ratings on four parameters, for Group A and Group B

Parameters	Group A		
	30 dB HL	45 dB HL	65 dB HL
Loudness	40	40	1280
Clarity	1280	1280 *	1280
Naturalness	1280 *	1280 *	1280 *
Overall Impression	1280 *	1280	1280
Parameters	**Group B**		
	30 dB HL	45 dB HL	65 dB HL
Loudness	40	40	640
Clarity	1280 *	1280 *	1280 *
Naturalness	1280 *	1280 *	1280 *
Overall Impression	1280 *	1280 *	1280

Table 7 summarizes the effect of three release time on SIS for Group A and Group B. As described in the above Table 7, at each input level, the release time which provides better ratings compared to other two release times are mentioned. Asterisks mark indicates that the particular release time is significantly better than the rest of the release times.

Loudness: Two-way repeated measures ANOVA did not reveal a significant interaction between the conditions and groups. The results did not reveal difference across nine conditions for both Group A and Group B (Table 7). As depicted in Table 5 and Table 6, the mean ratings are slightly higher at shorter (40 ms) release time. Although the effect is not seen at all conditions, this is in consonance with study by Neuman, Bakke, Hellman, and Levitt (1998). At shorter release times, the audibility will be maintained as the speech signal will be released from compression very quickly and the gain is applied for weaker part of the signal. Hence, the perceived loudness at short release times is due to the factor of audibility. This explanation is in agreement with that given by Moore (1996).

Clarity: Two-way repeated measure ANOVA did not reveal significant interaction between the conditions and groups. The results revealed a significant main effect across nine conditions [$F(8, 144) = 44.30$: $p<0.05$]. The effect is not significant in most of the conditions. But the mean ratings of clarity (Table 5 & Table 6)increased as the release time increased at all input levels, for both Group A and B. This is in agreement with the study by Neuman et al. (1998). At shorter release times, the gain fluctuates very rapidly giving the sensation of pumping of sounds (Wang, 2001; Moore, 1996). Hence, at longer release times the clarity of the speech sounds will be preserved to some extent in comparison with short release times.

Naturalness: The mean ratings of naturalness (Tables 5 & 6) increased as the release time increased at all input levels, for both Group A and B. Two-way repeated measure ANOVA did not reveal significant interaction between group and conditions. The results revealed significant difference across nine conditions [$F (8, 144) = 41.09$: $p<0.05$]. Bonferroni's multiple comparison (Table 9) revealed significant difference between 40 and 1280 ms, at 30 dBHL and 45dBHL; This was also seen between 40 and 640 ms, 40 and 1280 release time conditions, at 65 dBHL. The effect of release times on naturalness is the same as that of the effects of clarity. The results are in agreement with the study by Neuman et al. (1998). Since the distortions are controlled to some extent by using longer release times, the speech might sound pleasant and more natural.

Overall impression: As depicted in Tables 5, the mean ratings of overall impression increased as the release time was made longer, across the nine conditions, for both Group A and Group B. Two-way repeated measure ANOVA did not reveal significant interaction between groups and conditions. The results revealed significant difference across nine conditions [$F(8, 144) = 68.4$, $p<0.05$]. Bonferroni's multiple comparison (Table 9) revealed significantly higher scores at 1280 ms compared to 40 and 640 ms for most of the conditions. The results are in agreement with the study by Neuman et al. (1998). Due to the positive effects revealed by the longer release times on clarity, naturalness and other parameters, the overall impression will also be maintained at a higher rating for longer release times.

Effect of degree of hearing loss on quality rating: The Tables 5 and 6 depict the mean and SD of perceptual quality ratings of four parameters, across three release times and across three input levels, for Group A and B. In general, the perceptual ratings of quality are lower for Group B, though not significant, due to reduced dynamic range. The amount of compression taking place will be high and hence more temporal and spectral distortion in Group B. Hence, the quality ratings of the parameters are lesser for Group B.

Conclusions

From the present study it can be inferred that, 40 ms short release time induces significantly more temporal and spectral distortions when compared to 640 ms and 1280 ms release times. With respect to SIS and perceptual quality ratings, the performance was poorer with 40 ms short release time than when compared with 640 ms and 1280 ms release times. For both the acoustic and behavioural measures, as the input level increases or as the dynamic range reduces, the temporal envelope and spectral distortions also increases; and longer release times are required to offset the effects of reduced temporal envelope and spectral distortions.

The information from the present study helps the audiologists to gain knowledge as to how the compression release time can have various effects on temporal and spectral distortions. The current study helps to choose appropriate release time depending on the knowledge on dynamic range of the individuals with hearing impairment. The present study also helps to set appropriate release time when there is interaction between input level and compression ratio. Since the current study was carried out using hearing aid on real ear, it throws light on the realistic amount of distortions.

References

Bentler, R. A., & Nelson, J. A. (1997). Assessing release time options in a two-channel ACC hearing aid. *American Journal of Audiology, 6*(1), 43-51.

Dillon, H. (1996). Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? *Ear and Hearing, 17,* 287-307.

Fortune, T. W., Woodruff, B. D., & Preves, D. A. (1994). A new technique for quantifying temporal envelope contrasts. *Ear and Hearing, 15,* 93-99.

Freyman, R. L., Nerbonne, G. P., & Cote, H. A. (1991). Effect of consonant-vowel ratio modification on amplitude envelope cues for consonant recognition. *Journal of Speech and Hearing Research, 34,* 415-426.

Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear and Hearing, 23,* 369-380.

Henning, W. L. R., & Bentler, A, R (2008). The Effects of Hearing Aid Compression Parameters on the Short-Term Dynamic Range of Continuous Speech. *Journal of Speech, Language, and Hearing Research, 51,* 471-484 .

Hickson, L. M. H. (1994). Compression amplification in hearing aids. *American Journal of Audiology, 3*(3), 51-65.

Jenstad, L. M., & Souza, P. E. (2005). Quantifying the Effect of Compression Hearing Aid Release Time on Speech Acoustics and Intelligibility. *Journal of Speech Language & Hearing Research, 48*(3),651-67.

Jeon, Y., & Lee, S. (2008). *A strategy of contrast enhancement considering acoustic masking effect.* Paper presented at, The 2nd technical conference on rehabilitation engineering and assistive technology society of Korea.

Kuk, F. K. (1996). Theoretical and practical considerations in compression hearing aids. *Trends in Amplification, 1,* 5- 39.

Loizou, P. (1998). *Manual on COLEA- A MATLAB software tool for speech analysis.* http://ecs.utdallas.edu/loizou/speech/colea.htm.

Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing, 17,* 133-161.

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1996). Effect of release time in compression hearing aids: Paired-comparison judgments of quality. *Journal of Acoustic Society of America, 98,* 3182-3187.

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *Journal of the Acoustical Society of America, 103*(5), 2273-2281.

Novick, M. L., Bentler, R. A., Dittberner, A., & Flamme, G. A. (2001). Effects of release time and directionality on unilateral and bilateral hearing aid fittings in complex sound fields. *Journal of the American Academy of Audiology, 12*(10), 534-544.

Oxenham, A. J., & Bacon, P. S. (2003). Cochlear Compression: Perceptual Measures and Implications for Normal and Impaired Hearing. *Ear and hearing*, *24*(5), 352-356.

Ramkrishna, B. S., Nair, K. K., Chiplunkar, V. N., Atal, B. S., Ramachandran, V., & Subramanian, R. (1962). Some aspects of the relative efficiencies of Indian languages. *A study from information theory point of view.* Indian Institute of Science: Bangalore.

Rao, R. B., Murthy, R. B. M., & Rao, S. K. (2011). Speech Enhancement using a Modified Apriori SNR and Adaptive Spectral Gain Control. *International Journal of Computer Applications, 12,* 0975 - 8887.

Savithri. S, R. (1989). Acoustic and psychological correlates of speech, *Journal of Acoustic Society of India, 17,* (3,4), 41- 48.

Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends in Amplification, 6,* 131-165.

Van Tasell, D. J. (1993). Hearing loss, speech, and hearing aids. *Journal of Speech and Hearing Research*, *36,* 228-244.

Vanaja, C. S. & Jayaram, M. (2006). *Sensitivity and specificity of different audiological tests in differential diagnosis of auditory disorders*. Developed in Department of Audiology, AIISH, Mysore.

Walker, G., & Dillon, H. (1982). *Compression in hearing aids: An analysis, a review and some recommendations (NAL Rep. No. 90).* Canberra, Australia: National Acoustic Laboratories

Wang, K. B. (2001). Some considerations regarding attack and release times in compression hearing aids. *Hearing Journal. 54* (10), 36-45.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonetically balanced word list in Kannada.* A test developed at the department of audiology, Departmental project, AIISH, Mysore.

Comparison of Rhythm Perception in Dancers and Musicians

[1]Arya Chand & [2]K. Rajalakshmi

Abstract

The present study compared the rhythm perception abilities in Carnatic musicians (vocalists) and Bharatanatyam dancers. Twenty Carnatic musicians and 20 Bharatanatyam dancers (aged 19-39 years), with experience of 5 years or more in the respective fields were recruited for the study. The testing was conducted in 2 parts - (i) Synchronization with the rhythm of a musical excerpt (ii) Perceptual judgment of whether an imposed click train follows the beat of a musical excerpt. For the first part of the testing the participants were asked to tap the perceived rhythm of a given musical excerpt of average 30s duration. For the second part of the testing, participants were made to perceptuallyÂ judge whether a click track superimposed on a musical excerpt follows the rhythm of the excerpt (on beat condition), or if the click tempo is faster or slower than the musical excerpt or the clicks are not on beat of the excerpt (off beat condition). One way MANOVA was carried out for the statistical analysis. Results of the statistical analysis showed that there is no significant difference between dancers and musicians in the rhythm perception skills. Results also showed that experience correlated positively with rhythm skills.

Keywords: *Rhythm perception, dancers, musicians, beat perception.*

Introduction

Rhythm is defined as an ordered recurrent alternation of strong and weak elements in the flow of sound and silence. The experience of rhythm involves movement, regularity, grouping, accentuation and differentiation (Handel, 1989). Rhythm perception and production is a basic skill which helps us to synchronize with music like tapping, clapping, dancing, playing musical instruments and synchronizing oneself with other performers. The perception of rhythm is a dynamic process and it involves the synchronization of external musical stimuli with internal rhythmic processes (Jones & Boltz, 1989). Anatomically, the correlates of rhythm perception are attributed to basal ganglia (Grahn, 2009) and also to Supplementary Motor Area (bilaterally) and also extending into regions of cingulate gyrus, basal ganglia (Geisler, Zaehle, Jancke & Meyer, 2008).

Two important factors in rhythm perception are meter and pulse or beat. Beat is a perceived pulse that marks equally spaced points in time. Cooper and Meyer (1960) define pulse as a series of regularly recurring, precisely equivalent psychological events that arise in response to a musical rhythm. It is because of beat perception that we are able to relate the onset of temporal intervals as multiple or subdivisions of the beat, thus perceiving related intervals and not as unrelated intervals which in turn help in rhythm perception and production (Drake & Gerard, 1989; Ross & Houtsma, 1994; Patel, Iversen, Repp & Chen, 2005). Perception of beats occurs when we synchronize the external stimuli with an internally generated temporal pattern. The temporal properties which the listeners use to generate the internal rhythm are not clearly known. One property that may be important for beat perception in rhythm is the presence of simple integer ratio relationships between intervals in a sequence (Essens, 1986; Sakai et al., 1999). Another factor that aids beat perception is the perception of accents. Accents cause a particular note to feel more prominent than its surrounding notes. Previous research has shown that a listener's attention is attracted to accented events (Jones & Boltz, 1989). In musical contexts, the accents are created by non-temporal cues such as pitch, volume, and timbre, yet even rhythms without these cues can induce listeners to feel a beat internally (Brochard, Abecasis, Drake, Potter & Ragot, 2003).

The terms meter and metrical structure refer to patterns of regularly recurring stronger and weaker pulses (Lerdahl & Jackendoff, 1983). Lerdahl and Jackendoff's (1983) system, the fundamental pulse periodicity (the rate at which one might spontaneously tap with a musical rhythm) would be notated as a single row of beats, and the pattern of strong and weak pulses as additional rows of beats at related frequencies. Metrical structure or meter is temporal pattern created by the simultaneous perception of beats by different temporal scales.

Most studies in rhythm perception have compared musicians and non-musicians. Though dancers are a population who also depend on rhythm perception for their performance only few studies have pondered into the rhythm perception abilities of dancers. Hence a comparative study of rhythm skills in musicians and dancers would throw light onto which group have superior skills. The Bharatanatyam dance form and Carnatic music share similar musical compositions and are exposed to the same rhythms and they are trained similarly. But while performing and practicing, rhythm is maintained in different manner by the two groups. The Carnatic musicians maintain rhythm by tapping with the

[1]Email: aryachand46@gmail.com,
[2]Professor of Audiology, Email: veenasrijaya@gmail.com

palm and Bharatanatyam dancers maintain practice by stamping it on their feet. Understanding the perception of rhythm in dancers and musicians will throw light on which mode of rhythm maintenance could be utilized for training rhythm to the speech and hearing impaired individuals.

Beat perception is one factor expected to have an effect on rhythm perception. But the role of beat perception in rhythm is not studied much. Thus investigating the role of beat perception in rhythm can help us understand as to how to impart rhythm training for individuals with poor rhythm skills.

Method

Participants

Twenty individuals trained in Carnatic music and 20 others trained in Bharatanatyam dance were recruited for the present study. Age range of the participants varied between 19 years and 29 years. The participants were otologically and audiologically normal with hearing thresholds not exceeding 20dB at any frequencies between 250 to 8000Hz. The participants were also ruled out of any middle ear infections and any history of neurological disorders. All the recruited participants had received 5 years or above 5 years training in music or dance.

Instrumentation

A calibrated audiometer with TDH 39 headphones was used for Air Conduction testing and radio ear B-71 was used for bone conduction testing. GSI Tympstar was used for ruling out middle ear infections. Audacity software was used for generating the test stimuli and for recording the responses.

Procedure

The testing consisted of 2 parts; Synchronizing with the rhythm of a musical piece and Perceptual judgment of imposed rhythms on a musical piece.

Test 1: Synchronizing with the rhythm of a musical piece

Each participant was given 5 stimuli, (Carnatic instrumental pieces) one at a time. The musical excerpt had duration in the range of 25 to 30sec. The participants were made to listen to the stimuli twice and were asked to tap and report about the perceived rhythm of the piece. The participant's taps were recorded during the recording trail (3^{rd} trail) and were mixed with the original composition using the Audacity software. Information on the participants' tapping and about the meter of the composition as perceived was noted. The taps were compared with the original rhythm of musical excerpt and were also checked for the presence of any tapping errors in terms of phase or tempo. The participants were also asked about the familiarity of the musical composition.

The different rhythms selected were such that they formed the frequently taught and most common rhythms in both Carnatic music as well as Bharatanatyam dance.

Test 2: Perceptual judgment of imposed rhythms on a musical piece

For this part of the experiment, the participants were presented with a musical composition upon which a click train had been imposed. The click trains are superimposed in 3 conditions; hence there are three variations of the same composition. The three conditions are on beat and off beat (tempo error) and off beat (phase error).

On beat condition: For generation of this condition, the click train was generated with the same tempo as that of the compositions. The original tempo of the song was calculated by computing the beats per second (BPS). BPS was calculated by tapping to the song and calculating the number of taps per second. A click train generated with the same BPS was aligned to the composition in such a way that, the first click coincided with the original beats of the composition. Thus, the clicks fell exactly on the points where we expect the composition's beats to fall. A musically trained listener was made to judge whether the tempo was matching the composition's tempo.

Off beat condition (tempo error): For this condition the click train was generated in such a way that tempo of the click train was either greater or lesser than the actual tempo of the composition. For this condition, the original tempo of the song (BPS) was found out and was aligned with the composition for checking matching of the tempo. Once both the tempo matched, then the particular click train was removed and another click train was generated with a tempo which is 20% lesser than (slower tempo) or 20% faster than (faster tempo) that of the original tempo of the composition. This click train was aligned to the composition in such a way that, the first click of the composition falls on beat with the beat of the composition, but the tempo being different from the composition's original tempo.

Off beat condition (phase error): For generation of this condition, the click train with the same BPS as that of the song was generated, but the click train was aligned in such a way that, the first click fell either before or after the point where the actual beat of the song falls such that every time the clicks fell either before or after the intended beat.

The above mentioned three conditions were randomly mixed and the participants were required to report whether the click train imposed on a musical piece (Carnatic composition) follows the rhythm of that particular composition. The participants were exposed to 3 iterations of a stimuli such that in one falls in on-beat, in one, there is a off-beat in terms of phase error (early or late) and in the third one, there is a off-beat in terms of tempo error (slow or fast). Each subject was exposed to 5 stimuli, twice during testing. The participants were asked to report whether the clicks were falling on-beat with the composition or off beat with the composition. If the clicks are falling off -beat, the participants were asked to report whether there was an error in terms phase or tempo. If the error was reported to be phase, then participants had to report whether the clicks were early or late with reference to the beat of the composition and if the error was reported to be in tempo, then the participants had to report whether the beats were faster or slower when compared with the original composition. The participants were also asked regarding the years of experience with music or dance and familiarity with the musical compositions included in the study.

Scoring

Test one: Test one consisted of 5 compositions. For each composition, a maximum score of 3 was given, which was consisted of the scores for the three domains tested. The three domains tested are: Identification of the rhythm of a composition, synchrony of the participant's tapping phase with the phase of the composition and synchrony of the participant's tapping tempo with the tempo of the composition. A correct response in each domain was awarded 1 point. Thus, each composition gets a maximum of 3 points. Hence, the maximum total score in the Test one is 15.

Test Two: The stimuli in Test two was composed of 5 compositions each iterated 3 times (3 different conditions), thus making a total of 15 presentations of the stimuli. Identification of offbeat condition was awarded with 0.5 point. A score of 1 was awarded when the participant identified the exact offbeat condition in terms of whether it is a tempo error or phase error. The maximum score of 2 was awarded when (i) participant correctly identified whether in tempo error i.e. The click tempo as fast or slow than the composition's beats or (ii) participant correctly identified whether in phase error, the clicks were early or delayed with reference to the original composition. Hence the maximum score for each stimulus is 2. Thus, a total score of 30 is the maximum score in Test 2.

Results and Discussion

The data from 40 participants (20 dancers and 20 musicians) was subjected to statistical analysis. The data was tabulated and analyzed using SPSS (17.0). In order to compare the rhythm skills in dancers and musicians, a one way MANOVA was carried out between both the groups for the test scores obtained in the two tests, another one way MANOVA to compare the skills across the domains tested in each test, and correlation of various parameters of the two tests and correlation of the rhythm skills and experience of the participants.

Comparison of Scores Obtained in each domain tested in Test 1 between Musicians and Dancers (Musicians: Group A, Dancers: Group B)

The main aim of the study was to compare the rhythm skills of dancers and musicians. Thus the scores obtained for each domain in Test one was compared across the two groups using a one way MANOVA. The domains tested were (i) identification of the rhythm (ii) tempo synchrony (iii) phase synchrony. The mean, standard deviation, F value, significance level are shown in Table 1.

The p values for identification domain is p= 0.596, for tempo synchrony domain is p= 0.771 and phase synchrony domain is 0.912. Hence the statistical analysis showed that there is no significant difference between the two groups across any of the domains tested in Test one.

The results of the statistical analysis imply that the rhythm perception abilities in dancers and musicians are comparable. Participants in both the groups could identify the rhythm; synchronize with the musical composition according to its phase and tempo to the same extent. The reason for getting no significant difference between the two groups in the test can be attributed to factors like similarity in training imparted to both the groups, the rhythms selected for testing. The two groups considered under the study i.e. dancers and musicians would have both been exposed to the same kind of music during learning i.e., Carnatic Music since the rhythms and compositions used in the Carnatic music are commonly used in Bharatanatyam dance training. Also the rhythms used in the tests were common to both Carnatic music and Bharatanatyam dancers.

Table 1: Mean, Standard Deviation, F value, Significance for the two groups for the Test 1

Domains tested	Group	Mean	SD	F value	p
Identification	A	3.70	1.52	0.28	0.59
	B	3.45	1.43		
Tempo synchrony	A	4.25	1.06	0.08	0.77
	B	4.15	1.08		
Phase synchrony	A	3.90	1.48	0.01	0.91
	B	3.85	1.34		

Comparison of the Scores Obtained for the Identification of the Three Conditions in Test Two between Musicians and Dancers

Further to compare the rhythm skills under Test two, a one way MANOVA was carried out. The two groups were compared for their ability to identify 3 conditions of imposed beats on a musical composition namely- on-beat condition, off beat-tempo error condition and off beat-phase error condition. The mean, standard deviation, F value, are given in Table 2. Hence, the statistical analysis has shown that there is no significant difference between the 2 groups for identification of different conditions in the Test Two.

The results of the statistical analysis showed that both dancers and musicians have similar abilities in perception of beats. The perception was similar for perception of on-beat condition, off beat phase error condition and off beat tempo error condition. Previous study on general population (untrained in rhythm) by Iversen and Patel (2008) had indicated that on-beat conditions were identified more correctly than the tempo error or phase error condition. Such a pattern is not observed in musicians and dancers i.e. they identified off beat condition and on beat conditions easily. This could be attributed to the training effect and similarity in training rhythm in both the groups. The study of Iversen and Patel (2008) was carried out on general population who received no training in rhythm, hence the authors could not find out the similarity in perception of rhythm in their participants.

Comparison of the Total Scores for Both the test between the 2 Groups

For comparing the rhythm skill for the two tests (total score) for the two groups, the total scores obtained for each test were calculated and were compared using one way MANOVA.

Results from the statistical tests revealed no significant difference between the 2 groups for the two tests. The reasons for the non-significant difference between the two groups are because of similar training and rhythms chosen for testing. The rhythms chosen were present

Table 2: Mean, Standard deviation, F-value, Significance for the two groups for Test Two

Conditions	Group	Mean	SD	F value	p
On beat	A	8.80	1.88	0.03	0.85
	B	8. 90	1.51		
Off beat tempo error	A	7. 25	1.91	2.50	0.12
	B	6.30	1.88		
Off beat Phase error	A	5.82	2.3	0.60	0.43
	B	5.25	2.2		

Table 3: Mean, Standard deviation, F value for total scores of the two tests for the two groups

Test	Group	Mean	SD	F value	p
Test 1	A	11.65	3.963	0.02	0.88
	B	11.45	3.720		
Test 2	A	21.97	5.861	0.75	0.38
	B	20.45	4.695		

in both Carnatic music and Bharatanatyam dance. The compositions were also selected in such a way that it is not very common, so as to remove prior knowledge of the rhythm. It was noted during the testing that for those participants who had been exposed to the compositions before i.e. those who were trained in that particular composition had better scores than for those participants for whom the compositions were novel.

Computation of the Correlation between the Two Tests)

The second aim of the study was to find whether, perception of beat influences the synchronizing to a rhythm. To study this, correlation between the two tests used in the study was carried out. Correlation was found by considering the total 40 participants as a single group. The results of the analysis showed significant correlation for the two tests [Pearson's correlation coefficient: 0 .690 at 0.05 level of significance]. Correlation results are shown in Table 4.

Previous work on general population by Iverson and Patel (2008) on the correlation of beat perception and synchrony with a composition's rhythm showed weak correlation (correlation coefficient = 0.38; p<0.03). But in the current study, there is a positive correlation between the two indicating the enhanced abilities in dancers and musicians, owing to training effects and better exposure and familiarity with the compositions.

For each domain, the influence of Test one over Test two was found out by finding out the correlation between domains tested in Test one with the corresponding condition in Test two. Thus off-beat phase error condition in Test two was checked for correlation with phase synchrony in Test one. Similarly off beat tempo error condition in Test two was checked for correlation with tempo synchrony domain in Test one. The analysis resulted in significant correlation between detection of tempo error and the tempo synchronization ability (Pearson's Correlation Coefficient: 0.383 at 0.05 level of significance) and significant correlation between detection of phase error and phase synchronization ability (Pearson's correlation coefficient: 0.571 at 0.01 level of significance). Correlation results are shown in Table 4.

Thus, the positive correlation results between domains of Test one and Test two indicated that correct detection

Table 4: Results of correlation

Variable one	Variable two	Correlation Coefficient
Test one	Test two	0.690 **
Tempo error	Tempo synchrony	0.383*
Phase error	Phase synchrony	0.571**
Experience	Test one	0.679**
Experience	Identification of rhythm	0.710**

** Correlation is significant at the 0.05 level (2-tailed).*

of an off-beat tempo error aids in better synchrony with the tempo and a correct detection of an off-best phase error aids in better synchrony with the phase.

The third aim of the study was to find whether experience plays a role in rhythm perception. Thus correlation of the participants' test scores and the subject's experience was calculated using Pearson's Correlation coefficient. Results of the analysis have showed that experience was positively correlated significantly with test one scores [Pearson's correlation coefficient: 0.679 at 0.05 level of significance]. Correlation was calculated between identification scores in test one and experience of the participants. Results of the analysis, showed significant correlation between the two [Pearson's correlation coefficient: 0.710 at 0.01 level of significance]. Correlation results are shown in Table 4.

Thus, results of statistical analysis have showed that experience plays an important role in rhythm perception. In both the groups, participants with greater years of experience got better scores in rhythm. Thus experience is a factor which affects rhythm perception.

Previous study by Batalha and Macara (2008) had also got similar results. Their study has shown that when compared to dance students, professional dancers have better rhythm perception abilities. This is owing to better experience and exposure to the rhythm. In the present study, both dancers and musicians showed a positive correlation with experience. Thus confirming the fact experience plays an important role in rhythm perception.

Conclusions

Thus, the results of the present study indicated that dancers and musicians perform similarly on the identification of rhythm, tempo, synchrony and phase synchrony. Also, the performance of dancers and musicians was similar in on beat, off beat tempo error condition and off beat phase error condition. Further, beat perception has positive correlation with identification and synchronization with the rhythm.

References

Brochard, R., Abecasis, D., Potter, D., Ragot, R., Drake, C. (2003).The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science , 14*(4), 362-6.

Cooper. G., & Meyer, L.B. (1960). *The rhythmic structure of music.* Chicago: The University of Chicago.

Drake. C., & Gerard, C. (1989). A psychological pulse train: how young children use their cognitive framework to structure simple rhythms. *Psychological Research, 51,* 16 -22.

Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception & Psychophysics, 40*, 69-73.

Geisler, E., Zaehle, T., Jancke, L., & Meyer, M.(2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive neuroscience, 20*(3), 541-52.

Grahn,J.A. (2009). The Role of the Basal Ganglia in Beat Perception. *Annals of the New York Academy of Sciences, 1169,* 35-45.

Handel, S. (1989) *Listening*. Cambridge, MA: MIT press.

Iversen, J. R., & Patel, A. D., (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10).* Sapporo, Japan.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459-491.

Macara, A., & Batalha, A. P. (2008). Self-Perception of the Body of Dancers in Performance. *Proceedings Business Bodies.* International Conference - Universities of Finland (pp.175-181). Theatre Academy. IADE, Helsinki.

Patel, A., Iversen ,J. R., Repp, B. H., & Chen, Y. (2005). The influence of modality and metricality on synchronization with beat. *Experimental brain research, 163,* 226-238.

Ross. J., & Houtsma. A.J. (1994). Discrimination of auditory temporal patterns. *Perception and Psychophysics, 56*, 19-26.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience, 19*, 10074-10081.

Comparison between Preferred Gain and Nal-NL1 Prescribed Gain Formulae in Naive Adult Hearing Aid Users

[1]Chandan, H. S. & [2]Vijaya Kumar Narne

Abstract

The aim of the present study was to find the deviation in gain parameters, at three different input levels (soft, moderate and loud level) between preferred and prescribed (NAL NL-1) fitting strategy in naive hearing aid users. The study included 33 participants in the age range of 30 to 80yrs, with mild to severe cochlear hearing loss and using digital BTE hearing aid. The results of the present study revealed, the gain preferences relative to NAL-NL1 is different for different degree of hearing loss. Individuals with mild to moderate hearing loss preferred 3-4 dB lower whereas individuals with moderately severe and severe hearing loss preferred 4-8 dB higher gain than that prescribed by NAL-NL1. The preferred gain differences relative to NAL-NL1 across frequencies revealed greater deviations in mid frequencies than in low and high frequencies and increased with increase in degree of hearing loss.To conclude, Indian population requires higher gain at mid frequencies compared to western population for individuals with moderately severe and severe hearing loss.

Keywords: *Preferred gain, prescribed gain, NAL-NL1 formula.*

Introduction

Cochlear hearing loss can vary in terms of degree and configuration. This necessitates tailor made fitting of the hearing aid for every client. Widely practiced approach in the clinics is to use a prescriptive procedure, to provide approximate target amplification. The prescriptive approach for hearing aid fitting is, one in which the amplification characteristics are calculated from some of the hearing characteristics of the individual. This is based on the assumption that certain amplification characteristics suit certain type, degree and configuration of hearing loss. The prescriptive methods were changed over the years due to advancement in technology, better understanding of hearing characteristics and other factors affecting hearing aid performance.

The prescriptive formulae, threshold based or supra threshold based, give the first approximation of gain required. Clinical experiences with prescriptive methods show that the methods cannot eliminate the need for individual adjustments i.e., fine tuning of hearing aid (Dillon, 2001). However, one should bear that fine tuning of gain settings in the hearing aids is performed on prescribed gain. The prescribed gain should be a good approximation to preferred gain, which reduces the trial and error done by the clinician and also saves time (Dillon, 2001).

The gain preferred by naive hearing aid users is lesser than that preferred by experienced users (Humes, Wilson, Barlow & Garner, 2002;Smeds, 2004). The new hearing aid users require lesser gain than that prescribed by NAL-RP (Byrne & Cotton, 1988; Cox & Alexander, 1992; Horwitz& Turner, 1997; Humes, Wilson, Barlow, Garner, 2002, Smeds, 2004). Many investigators from western countries (using English speaking individuals) compared NAL-NL1 formula in naive and experienced hearing aid users (Keidser et al, 2001; Keidser et al., 2006, Keidser et al., 2008; Humes et al., 2002). They observed that NAL-NL1 provides 3 to 6 dB higher gain than that preferred by cochlear hearing loss individuals.

Most of the above studies comparing preferred and prescribed gain were performed on western population. On contrary, Indian population, Mathur and Manjula (2008) showed opposite results, that is Indian populations prefer higher gain than that prescribed by NAL-NL1. However, the studies done by Mathur and Manjula (2008) considered overall gain from hearing aid software program. The difference that is noted between NAL-NL1 and preferred gain mayn't be appropriate because the target formula selected as NAL-NL1 in programming software gives a lesser gain of 10dB than the target gain noted from REIG values. So, further studies are needed to confirm the results of the previous study. There are only limited studies available on comparing preferred gain and prescriptive gain settings in naive hearing aid users in Indian context. Further, general opinion among the clinicians in India is that, majority of the clients prefer different gain settings than that prescribed by NAL-NL1. Hence, it becomes all the more important to compare the prescribed and preferred gain settings in new hearing aid listeners. These deviations can be studied using REIG and speech perception measures.

It is important to use REIG measurements because it provides the true gain in the ear canal. Aazh and Moore (2007) have demonstrated that, the currently available programming software provides an inappropriate gain in the ear canal than that prescribed by the prescriptive procedures. Hence, measuring REIG is an essential tool while fitting the hearing aid.

Many researchers have used speech perception mea-

[1]Email: chandanhs6@gmail.com,
[2]Lecturer in Audiology, Email: vijaynarne@gmail.com

sures to check for the acceptance of hearing aid gain characteristics. Some have used continuous discourse with noise (Keidser et al., 2005) and a few others have used speech recognition threshold (Moore, Alcantara & Marriage, 2001). The mentioned studies demonstrate that scores were different between preferred and prescribed condition. Prescriptive formula provides maximum emphasis to speech; it thus becomes an important tool to study the difference in speech perception using preferred and prescribed condition.

The aim of the present study is to find the deviation in gain parameters, at three different input levels (soft, moderate and loud level) between preferred and prescribed (NAL NL-1) fitting strategy in naive hearing aid users.

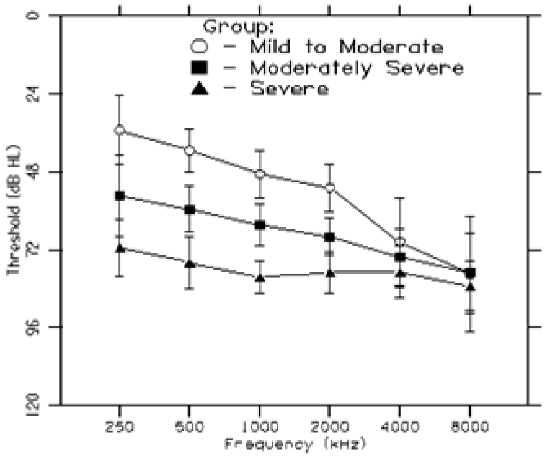


*Figure 1: Mean and standard deviation of pure-tone thresholds for 3 individual **group**.*

Method

Participants

Thirty three participants (n=33 ears), in the age range of 30 to 80yrs (mean age of 59.2yrs.), with mild to severe cochlear hearing loss and using digital BTE hearing aid were participated in the present study. The participants were native speakers of Kannada (A Dravidian language spoken in a southern state of India), having post-lingual onset of hearing loss and were naive hearing aid users (with duration of Hearing aid use not more than 3 months). The Pure-Tone Average (PTA) ranged from 36.6 dB to 85 dB. The participants were divided into 3 groups namely, Group I, Group II and Group III based on the degree of hearing loss. Group I included 13, Group II and Group III included 10 participants with Mild to Moderate, Moderately severe and severe hearing loss respectively. It was ascertained from a structured interview that none of these participants had any history of neurologic disorders. The mean and standard deviation of pure-tone thresholds at octave frequencies for all the 3 individual groups is plotted in Figure 1.

The experiments were conducted in two phases. In the phase 1 hearing aid was programmed to NAL-NL1 settings, which was followed by measurements of speech Identification Scores. In phase 2, hearing aid was programmed to the participant's preferred setting, which was followed by measurements of speech Identification Scores.

Phase 1

Initially the hearing aid was programmed accordingto the gain parameters prescribed by NAL-NL1 fitting formula as given by hearing aid fitting software. 'First fit' settings were obtained by using the participant's hearing thresholds and selecting NAL-NL1 prescriptive formula. It has been noted by number of researchers that the hearing aid programming software provides an inappropriate gain than that prescribed by the NAL-NL1 prescriptive procedure (Aazh & Moore,2007). So, REIG was performed to attain appropriate gain parameters of NAL-NL1 by matching REIG value to gain curve generated by FONIX 7000 hearing aid analyzer for NAL-NL1 target. However, it was not possible to achieve the perfect match to the target in all the participants. The relationship between the achieved and NAL-NL1 prescribed 4 frequency averages (4FA) i.e. 500, 1000, 2000 and 4000 Hz for all the participants is given in Figure 2. From the figure it can be inferred that there wasn't much difference between achieved and target REIG values. The mean difference between the achieved and target REIG values for 4FA is 3 dB. The average fit was closer to target at mid frequencies than in low and high frequencies. After approximating the REIG values to FONIX target curve, these values were considered as the NAL-NL1 gain prescribed. Following this Speech Identification Score (%) was measured.

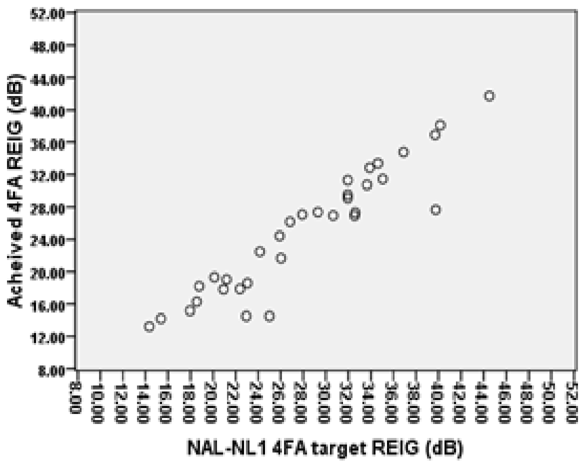


Figure 2: The relationship between the achieved and NAL-NL1 prescribed 4 frequency averages (4FA) for all the participants.

Phase 2

The hearing aid was programmed as per the preference of the clients. Initially, the automatic fine tuning of the hearing aid was carried out using 'feature activation' or 'Fitting Assistant' designed specifically for hearing aid fitment. Later, 'manual fine tuning' was done by narrating a passage in Kannada at moderate and loud intensity levels. With this, all or few of the hearing aid parameters (listed below) were manipulated as preferred by the participants. The parameters that varied were Overall gain, Gain at individual Frequency Bands, Enhanced Bass Boost and global Compression. The adjustment of overall gain and individual frequency bands gain was performed at 65 dB input level only. It was noted that the gain of 50 dB and 80 dB were also varied while fine tuning the gain for 65 dB input level. Thus, the gain at 50 dB and 80 dB was not changed individually. After, fine tuning of the hearing aid for preferred hearing aid settings, REIG and speech identification scores (%) are measured again.

Speech Identification Scores: The open set Speech Identification Scores in quiet were obtained through monitored live voice presentation. Word lists for adults developed by Yathiraj and Vijayalakshmi (2005) was used to obtain the speech identification scores. This test material consisted of 4 phonemically balanced word lists with 25 words each.

The participants were seated comfortably in a double-walled, acoustically treated room. The speech stimuli were presented through the loudspeaker (C 115 Martin Audio) of the audiometer kept at a distance of one meter at 0Â° azimuth. Speech stimuli were presented at 40 dBHL. None of the lists was repeated for any of the listeners, as there were four lists. The order of presentation of conditions was randomized across the listeners. Listeners were instructed to repeat the speech token heard. The speech recognition scores were calculated by counting the number of words correctly repeated.

Real Ear Insertion Gain (REIG) : REIG, as defined by ANSI (1997), is the difference in decibels as a function of frequency between the real ear aided gain (REAG) and real ear unaided gain (REUG), obtained with the at same measurement point and similar sound field conditions. Before the REUG is measured, levelling of the probe system of the hearing aid analyser instrument was done using the reference microphone placed above the ear to ensure the smooth frequency output from the hearing aid analyser.

Measurement of REIG: The participants were seated at 1 foot distance and at 45 degree azimuth from the loudspeaker of real ear analyser. Real Ear Unaided Gain (REUG) was measured for the subjects without wearing the hearing aid by using Digispeech as the stimuli at 65dBSPL as the input. To ensure proper insertion depth of the probe tube, the probe tube was placed in the ear canal, so that the tube will rest along the bottom of the canal part of the ear mould, with the tube extending at least 5 mm past the ear mould. The stimulus was presented and the output was represented in the form of graph on screen and once the graph on screen is stabilized for more than 10 seconds, the input was stopped. For measuring REAG, the hearing aid was placed into the participant's ear while holding the probe tube so that its position in the ear canal is not disturbed. Then, hearing aid is turned on for measuring REAG. The probe tube microphone measures the dB SPL in the ear canal as delivered by the hearing aid. The REAG was displayed as a curve with frequency (Hz) versus Intensity (dB).The real ear analyser automatically displayed the REIG across frequencies. This was done by the instrument by subtracting REUG from REAG. The values of REIG were noted across 250 Hz, 500 Hz, 700 Hz, 1k Hz,1.5k Hz,2k Hz,3k Hz, 4k Hz, and 6k Hz for each participant. The REIG was also calculated at 3 different input levels i.e., 50, 65 and 80 dB SPL.

Results

The present study was carried to find the difference in gain between preferred and prescribed (NAL NL-1) strategies in naive hearing aid users at three different input levels (50, 65 & 80dB). The REIG data at three input levels (50, 65 and 80 dB levels) were collected, tabulated and subjected to data analysis. Statistical analyses were carried out using SPSS Statistics Package (version 17).

Comparison of preferred and prescribed Real Ear Insertion Gain (REIG) values

Majority of previous studies have compared results for overall gain (4FA), LFA (250, 500 and 1000 Hz) and HFA (2000, 3000 and 4000 Hz) at 65 dB input level for different degree of hearing loss. Hence for ease of comparison only 65 dB input level is considered for the analysis of overall gain (4FA), LFA & HFA.

Relationship between preferred 4FA gain relative to NAL-NL1 and the pure tone average (PTA): Figure 3 gives the relationship between the preferred 4FA gain relative to NAL-NL1 and the pure tone average (PTA) for all the participants. It can be inferred from the Figure 3 that with increase in hearing loss, there was increment in gain deviation i.e. the gain deviation was higher for greater degree of hearing loss.

From the REIG measurements, gain deviation of preferred from the NAL-NL1 for a 65 dBSPL input was calculated in terms of overall gain (4FA), LFA and HFA for individual groups. The mean and standard deviation of this for three individual groups is given in Figure 4. It can be noted from the figure that the average gain preferred is around 2 dB lower for group 1, whereas group

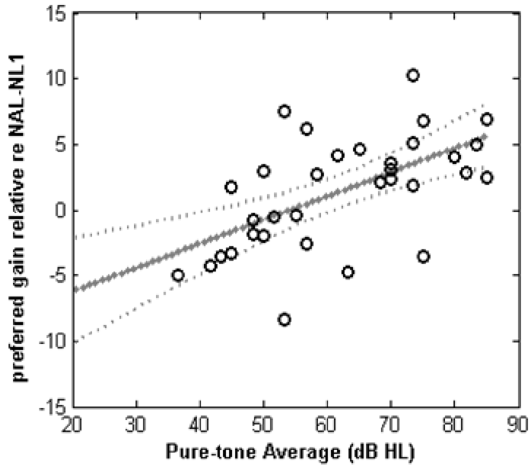


Figure 3: Shows the relationship between the preferred 4FA gain relative to NAL-NL1 and the pure tone average (PTA). The dotted lines show the regression line and 95% confidence bands.

2 and group 3 preferred is 2 to 4 dB higher gain than that prescribed by NAL-NL1. Mixed ANOVA was performed to compare the conditions (LFA, HFA & Overall (4FA) across the three groups. Analysis showed no significant main effect of conditions, indicating that mean difference did not reach significance between conditions [$F_{(2,30)}$ =0.7, P= 0.40]. But there was significant main effect of groups [$F_{(2,30)}$ =5.09, P<0.05]. Bonferonni's Post hoc analysis revealed group 1 is significantly different from group 2 and group 3 (p <0.05). But mean difference between group 2 and group 3 did not reach significance.

Difference in preferred gain and prescribed gain across frequencies: Further to know the difference in REIG between preferred and that prescribed by NAL-NL1 at each frequency across three groups for different input levels a separate analysis was done.

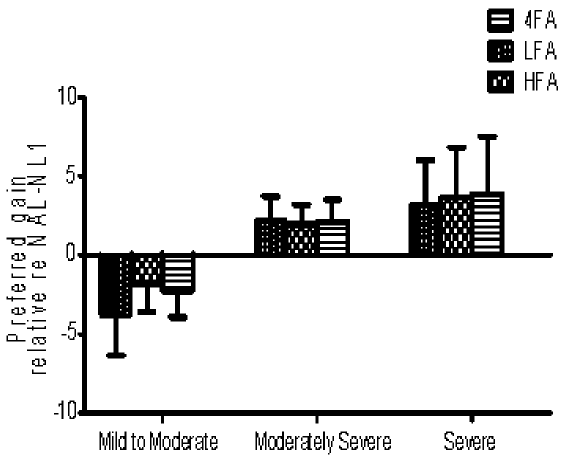


Figure 4: The comparisons of REIG values at 3 input levels for group 1.

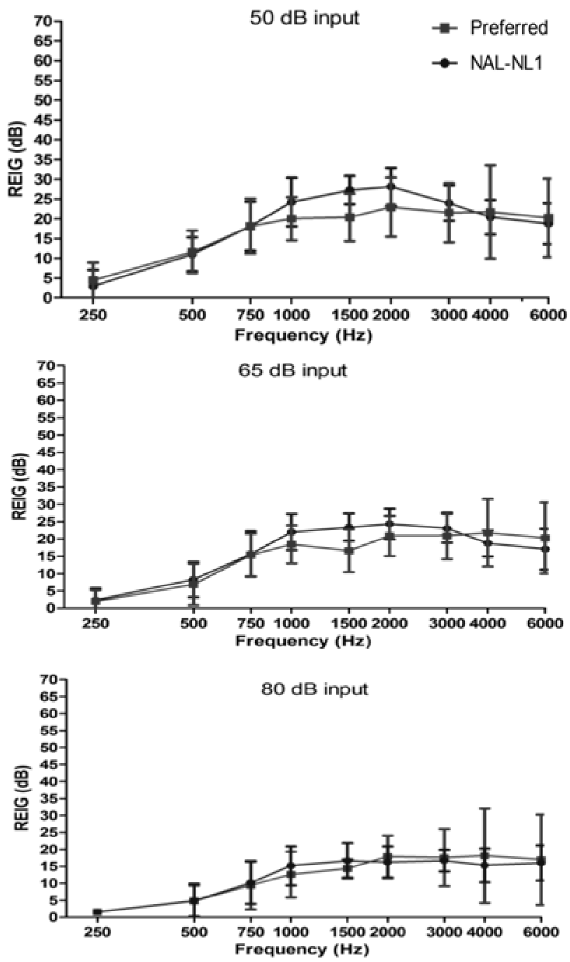


Figure 5: comparisons of REIG values at 3 input levels for group 2.

Group 1-Mild to Moderate HL

Figure 5 shows the preferred and prescribed REIG data across frequencies for group 1. It can be noted from the figure that for 50 dB and 65dB input levels, individuals with Mild to Moderate hearing loss preferred a gain 5 to 8 dB lower than that prescribed by NAL-NL1 at mid frequencies. However, at 80 dB input level, preferred and prescribed gains were almost similar. Further, it was also noted that there was no difference in mean REIG values at low and high frequencies for all 3 input levels.

Repeated measure ANOVA was performed to compare the gain between conditions (preferred and NAL-NL1) at three input levels (50, 65 & 80dB) across the frequencies. Analysis showed significant main effect frequency [F (2.9, 106.7) =133.5, p<0.01] and level [F (2, 36,) =7.5, p<0.01] but no significant main effect of condition [F (1, 36) =3.59, p=0.06]. Interaction analysis-revealed significant interaction between frequency and condition [F (3.2, 36) =8.01, p<0.01], but other two way and three way interactions were not significant. Following this a Paired sample't' test was performed to assess at which frequencies, difference between conditions reaches significance for three different input levels separately. The results of 't' test, degrees of freedom and level of significance are depicted in Table 1. It can be noted from the Table 1 that for 50 and 65dB input level there was significant difference between preferred and prescribed conditions at 1, 1.5 and 2 kHz only. But, there was no significant difference across frequencies for 80dB input.

Table 1: Shows the 't' values and level of significance for three input level across frequencies for group 1

	Input Level		
	50dB	65dB	80dB
Frequency(Hz)	t value	t value	t value
250	1.08	-0.65	-0.06
500	0.44	-1.51	-0.08
750	0.0	-0.24	-0.54
1000	-2.97*	-3.3*	-1.78
1500	-4.3*	-4.7*	-1.83
2000	-3.2*	-2.4*	1.19
3000	-1.18	-1.42	0.44
4000	0.42	1.28	1.23
6000	0.62	1.46	0.45

**p<0.05, Note: df was 12 for all't' values*

Group 2- Moderately Severe Hearing loss

Figure 6 gives the preferred and prescribed REIG data across frequencies for group 2. It can be noted from the figure that for the input level of 50 dB, 2-3 dB higher gain is preferred at low and mid frequencies than that prescribed by NAL-NL1. Whereas for input levels of 65dB and 80 dB, gain preferred is 5-6 dB higher than NAL-NL1 at mid frequencies.

Table 2: 't' values and level of significance for three input level across frequencies for group 2

	Input Level		
	50dB	65dB	80dB
Frequency(Hz)	t value	t value	t value
250	1.66	1.74	2.92*
500	0.90	1.41	2.08
750	2.38*	2.68*	3.27*
1000	1.23	2.38*	2.73*
1500	1.62	2.32*	2.51*
2000	1.09	2.95*	2.40*
3000	0.13	0.75	1.60
4000	0.14	1.45	1.08
6000	-0.08	3.20*	0.53

**p<0.05, Note: df was 12 for all't' values*

Repeated Measure ANOVA was performed to compare the gain between conditions (preferred and NAL-NL1) at three input levels (50, 65 & 80dB) across the frequencies. Analysis showed significant main effect of frequency [F (8, 20) =202.6, p<0.01], condition [F (1, 27) =13.15, p<0.01] and level [F (2, 27,) =6.4, p<0.05].Interaction analysis revealed significant interaction between frequency and condition [F (3.04, 82.2) =3.05, p<0.05[, but other two way and three way interactions were not significant. Following this a Paired sample't' test was performed to assess at which frequencies difference between conditions reached significance for three different input levels. The results of 't' test, degrees of freedom and level of significance are depicted in table 2. It can be noted from the Table 2 that there was significant difference across 0.75, 1, 1.5, 2, and 6 kHz for 65dB input level and 0.25, 0.75, 1, 1.5, and 2 kHz for 80dB input level between preferred and prescribed conditions. However, for 50 input levels there was significant difference at 0.75 kHz only.

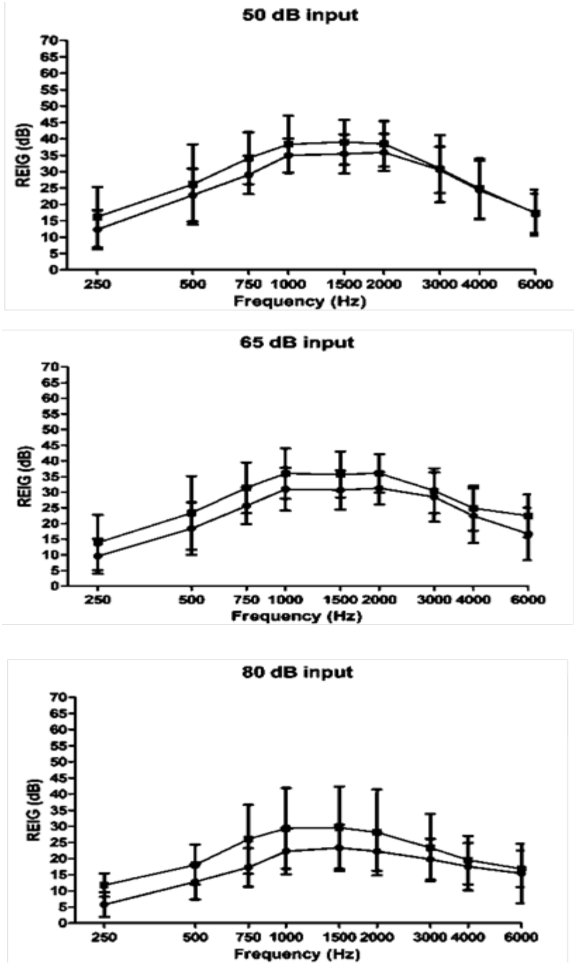


Figure 6: comparisons of REIG values at 3 input levels for group 2.

Group 3-Severe Hearing Loss

Figure 7 gives the preferred and prescribed REIG data across frequencies for group 3. It can be noted from the figure that the preferred gain was higher than that pre-

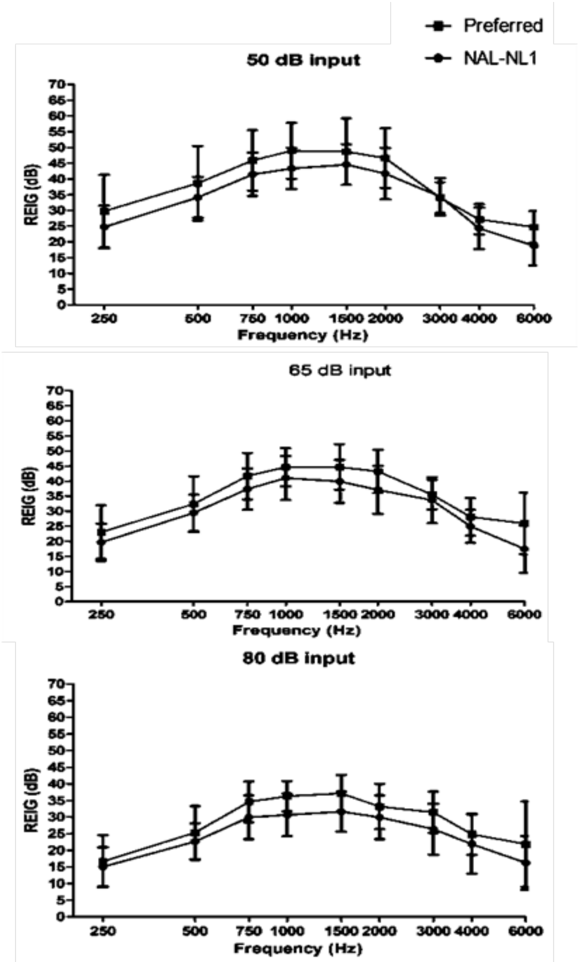


Figure 7: The comparisons of REIG values at 3 input levels for group 1.

scribed by NAL-NL1 for all the 3 different inputs. For 50dB and 65dB input, 5-6 dB higher gain is preferred at mid frequencies but, for 80dB input the gain of 7-8 dB higher is preferred at mid frequencies. Repeated Measure ANOVA was performed to compare the gain between conditions (preferred and NAL-NL1) at three input levels (50, 65 &80dB) across the frequencies. Analysis showed significant main effect frequency [F (1.6, 45.08) =72.3, $p<0.01$], level [F (2,27) =10.4, $p<0.01$] and condition [F (1, 27) =42.56, $p<0.01$]. Interaction analysis revealed no significant interaction between any conditions. Following this a Paired sample't' test was performed to assess at which frequencies difference between conditions reaches significance for three different input levels.

The results of 't' test ,degrees of freedom and level of significance are depicted in table 3.It can be noted from the table 3 that for 50 input levels there was significant difference between preferred and prescribed conditions at 750, 1000, 2000 & 6000Hz. However, there was significant difference across 750, 1500, 2000, 4000, 6000 Hz for 65dB input and 750, 1000, 1500, 3000Hz for 80dB input.

Table 3: 't' values and level of significance for three input level across frequencies for group 3

	Input Level		
	50dB	65dB	80dB
Frequency(Hz)	t value	t value	t value
250	1.74	1.47	0.58
500	1.58	1.30	1.22
750	2.28*	4.36*	4.76*
1000	3.09*	2.01	4.33*
1500	2.15	2.72*	4.31*
2000	2.34*	3.96*	1.40
3000	-0.14	0.91	2.35*
4000	1.69	2.37*	1.31
6000	3.81*	3.9*	1.80

**p<0.05, Note: df was 12 for all't' values*

Comparisons of Aided Speech Identification Scores (%) between preferred and prescribed conditions

Figure 8 shows the comparison of Aided Speech Identification Scores (%) between prescribed and preferred settings across three groups. It can be noted from the figure that there is 5% increase in SIS in preferred gain than in NAL-NL1 settings in Moderately Severe and Severe Hearing Loss individuals. But,there was no difference noted in the mean of Mild to Moderate Hearing Loss group.

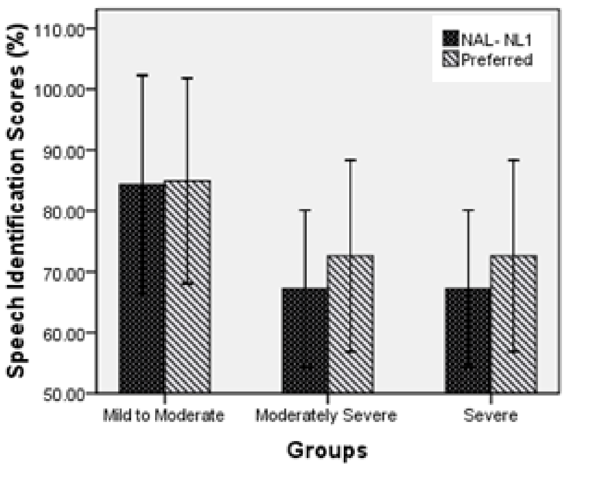


Figure 8: comparison of Aided Speech Identification Scores (%) between prescribed and preferred settings across three groups.

Repeated Measure ANOVA was performed to compare the Aided Speech Identification Scores (%) between conditions (preferred and NAL-NL1) across three groups. Analysis showed no significant main effect between conditions ($F_{(1,30.1)}$ =3.1, p=0.27) and groups ($F_{(2,40.2)}$ =0.75, p=0.4). Bonferonni's Post hoc analysis revealed there is no significant difference between groups.

Discussion

The aim of the present study was to find the deviation in gain parameters, at three different input levels (soft, moderate and loud levels) between preferred and prescribed (NAL NL-1) gain in naive hearing aid users. This was investigated by comparing REIG values and Speech Identification Scores (%) between preferred and NAL-NL1 settings.

Comparison of Preferred and Prescribed Gain

The present study compared REIG values between preferred and NAL-NL1 for different degrees of hearing loss. The results of overall gain showed individuals with Mild to Moderate hearing loss preferred 3 to 4 dB lower but Moderately Severe to Severe hearing loss preferred 4-8 dB higher gain than that prescribed by NAL-NL1. Mathur and Manjula (2008) compared the preferred and prescribed gain in naive hearing aid users. They reported individuals with Moderate and Moderately Severe hearing loss prefer 2 to 5 dB lower but severe hearing loss preferred 4 to 7 dB higher gain than that prescribed by NAL-NL1.

The results of the present study are in accordance with those reported by Mathur and Manjula(2008). However, they demonstrated individuals with Moderately Severe hearing loss preferred lower gain of 2 to 5 dB but in the present study higher gain of 3 to 6 dB was preferred. The precise reason for the difference is not known. Difference between studies may be due to methodological differences. One potential methodological reason could be in the present study REIG was matched to FONIX 7000 hearing aid analyzer target for NAL-NL1 but this was not done in Mathur and Manjula (2008).

In the contrary to the present study, Keidser et al., (2001; 2004; 2005;2008) reported that naive hearing aid users requires 2 to 6 dB lesser gain with reference to gain provided by NAL-NL1. The difference in gain preferred with reference to NAL-NL1 among these studies may be due to the subject population. Keidser et al. (2001), Keidser et al. (2004) Keidser et al. (2008) investigated on western population and showed that preferred gain is lower than that prescribed by NAL-NL1. In contrary, studies conducted in Indian population showed that gain preferred is higher than NAL-NL1 (Mathur & Manjula, 2008; Achaiah & Narne, 2011).

Achaiah and Narne (2011) reported on an average 10 dB higher gain is preferred compared to NAL-NL1 fitting formula in experienced hearing aid users. Higher difference noted between present study and Achaiah and Narne (2011) study may be because they have considered experienced hearing aid users. In addition to that they have not matched REIG values to FONIX 7000 Hearing Aid analyzer for prescribed NAL-NL1 gain settings. These results are in agreement with clinical observation made by majority of the clinicians in Indian population. The precise reason for needing a higher gain is not known. Probable reason could be that, as Studebaker and Sherbecoe (1993) reported that frequency importance functions vary widely across the languages and hearing aid prescriptive formulae were derived from the frequency importance function. Probably, the frequency importance functions for Indian languages are different which would have led to this difference.

The analysis was carried across three different frequency averages (LFA, HFA, and 4FA (overall gain) for three groups. It was noted that the overall gain preferred by Moderately Severe and Severe hearing loss subjects were 3-4 dB higher than that prescribed by NAL-NL1. Whereas, Mild to Moderate degree of hearing loss preferred -4 dB lesser gain than that prescribed by NAL-NL1. In order to understand which frequencies were showing the difference in REIG between preferred and prescribed by NAL-NL1, further analysis was carried out. The results of these analysis revealed the gain differences were noted only in mid frequencies. There was no consistent gain difference in low and high frequencies. These may be attributed to frequency importance function of Indian languages. In addition to the gain differences noted at 65 dB input level, the gain differences were also observed in 50 dB & 80 dB input levels. To our knowledge there were no studies that have compared the gain differences at 50 dB & 80 dB input level. One logical reason for gain difference noted was because gains at other input levels were also modulated by varying the gain at 65dB input level in the hearing aid programming software, which would have led to these differences.

Present study also analyzed the relationship between the pure tone average and preferred gain relative to NAL-NL1. It was noted that with the increase in degree of hearing loss, the preferred gain relative to NAL-NL1 increased. This finding are contrary to the findings of Keidser et al. (2008), who reported gain preferred relative to NAL-NL1 was lower with increase in hearing loss. These difference in findings noted in Keidser et al. (2008) study may be because majority of subjects considered were individuals with Mild to Moderate hearing loss but in the present study, the subjects were evenly distributed between different degree of hearing loss.

Speech Identification Scores

The Aided Speech Identification Scores (%) between prescribed and preferred settings across three groups were subjected to analysis. It is observed that there is only a 5% mean difference among the group in SIS (%) scores in group 2 and group 3. However, there is no significant differences observed between these groups, this could be attributed to less number of subjects and greater variability (as indicated by large standard deviation). Though there was no significant difference in preferred and NAL-NL1 gain condition on SIS (%), yet the individuals preferred different gain settings over that prescribed by NAL-NL1, for the overall enhancement in speech quality.

Conclusion

The major findings of the study indicated that the preferred gain differences relative to NAL-NL1 across frequencies was greater at mid frequencies than at low and high frequencies. The differences increased with increase in degree of hearing loss. This study also reflects on the importance of fine-tuning of hearing aids based on participant's preference and to develop a new prescriptive formula specifically for Indian population.

References

Aazh, H., & Moore, B. C. J. (2001).The Value of Routine Real Ear Measurement of the Gain of Digital Hearing Aids.*Journal of American Academy of Audiology*, *18*, 653-664.

Achaiah, M.A., & Narne, V. K. (2011). Comparison between outcomes using preferred gain and prescribed gain formulae in experienced adult hearing aid users. Unpublished Masters dissertation. University of Mysore, Mysore.

Mathur, A., & Manjula. P. (2008). Efficacy of NAL-NL1 Prescription for the First-Time-Hearing Aid users: A Follow up Study. Unpublished Masters Dissertation. University of Mysore, Mysore.

ANSI-S3.5. (1997).American national standard methods for the calculation of the speech intelligibility index. New York: American National Standards Institute.

Byrne, D., & Cotton, S. (1988). Evaluation of the National Acoustic Laboratories' new hearing-aid selection procedure. *Journal of Speech & Hearing Research*, *31,* 178-186.

Cox, R. M.,& Alexander, G. C. (1992). Maturation of hearing aid benefit: Objective and subjective measurements. *Ear and Hearing* , *13,* 131-141.

Dillon, H (2001). *Hearing aids*. Turramurra: Boomerang Press.

Horwitz, A. R., & Turner, C. W. (1997). The time course of hearing aid benefit. *Ear and Hearing*, *18*, 1-11.

Humes, L. E., & Christensen, L. (1999). Comparison of the aided performance and benefit provided by a linear and a two-channel wide dynamic range compression hearing aid.*Journal of Speech Language & Hearing Research, 42,* 65-79.

Humes, L. E., Wilson, D. L., Barlow, N. N., & Garner, C. (2002). Changes in hearing-aid benefit following one or two years of hearing-aid use by older adults.*Journal of Speech, Language & Hearing Research*, *46*, 137-145.

Keidser, G., & Grant, F. (2001). Comparing loudness normalization (IHAFF) with speech intelligibility maximization (NAL NL-1) when implemented in a two-channel device. *Ear and Hearing*, *22*, 501-515.

Keidser, G., Brew, C., Brewer, S., Dillon, H., Storey. L., & Grant, F. (2005). The preferred response slopes and two-channel compression ratios in twenty listening conditions by hearing-impaired and normal-hearing listeners and their relationship to the acoustic input. *International Journal of Audiology*, *44,* 656-670.

Keidser, G., & Dillon, H. (2006). What's new in prescriptive fittings Down Under? In Palmer, C.V., & Seewald, R. (Eds.), *Hearing Care for Adults*. pp. 133-142. Phonak AG, Stafa, Switzerland.

Keidser, G., O'Brien, A., Carter, L., McLelland, M., & Yeend, I. (2008). Variation in preferred gain with experience for hearing aid users. *International Journal of Audiology*, *47,* 621-635.

Smeds, K. (2004). Is normal or less than overall loudness preferred by first- time hearing aid users?.*Ear and Hearing*, *25,* 159-172.

Studebaker, G. A., & Sherbecoe, R. L. (1993). Frequency-importance functions for speech recognition. In G. A. Studebaker & I. Hochberg (eds.), *Acoustical factors affecting hearing aid performance* (2nd ed., pp. 185-204). Boston: Allyn and Baacon.

Yathiraj, A., & Vijayalakshmi.(2005). Auditory memory test. Test developed at the Department of Audiology, All India Institute of Speech and Hearing, Mysore.

Effect of Age on Spectral Distribution of Click and Toneburst Evoked Otoacoustic Emissions in Infants

[1]Chandni Mallik & [2]Mamatha N.M.

Abstract

Otoacoustic emission (OAE) is a sensitive tool used to assess the functional status of the outer hair cells (i.e., the cochlear component most vulnerable to many of the diseases and disorders that damage hearing) in infants with normal hearing and in different clinical population. The present study aimed to investigate the influence of age on click and toneburst evoked OAEs, to investigate the pattern of frequency shift with age, and to monitor the frequency specific maturational changes in the cochlea. The study was conducted on 40 infants (80 ears) with normal hearing sensitivity in the age range of 0 to 12 months, with subgroups as group I, II, III, IV. The results of the present study revealed that, the mean absolute amplitude of click and toneburst evoked OAEs across age groups increased as the age increased. Similarly, the mean SNR values of click evoked OAEs across age groups increased as the age increases. Whereas, the mean SNR values of toneburst evoked OAEs across age groups increased as the age increases, except for group I. Also, statistically significant difference for mean absolute amplitude was observed between CEOAEs and TBOAEs across all the frequencies. These findings can be used as a clinical tool for screening and diagnostic purposes in the pediatric population.

Keywords: *Otoacoustic emissions, click evoked otoacoustic emissions, toneburst evoked otoacoustic emissions.*

Introduction

Otoacoustic emissions (OAEs) are sounds that originate in the cochlea and propagate through the middle ear into the ear canal, where they can be measured using a sensitive microphone. Because of its relative simplicity, better sensitivity and objectivity of the technique, the OAEs are a promising means for monitoring cochlear function. OAEs are the property of healthy normal functioning cochlea, generated by active frequency selective, non-linear elements within the partition, the critical components being the outer hair cells (Kemp, 1988). The primary purpose of otoacoustic emission (OAE) tests is, to determine cochlear status, specifically the outer hair cell function. This information can be used to screen for hearing, particularly in neonates, infants, or individuals with developmental disabilities, partially estimate hearing sensitivity within a limited range, differentiate between the sensory and neural components of sensorineural hearing loss, and test for functional (feigned) hearing loss. The development of evoked OAEs has gained much interest because they are used as a valid and relatively quick test to assess cochlear integrity in the very youngest subjects.

Transient evoked OAEs (TEOAEs) are frequency dispersive responses following a brief acoustic stimulus such as click or tone burst (Kemp, 1978; Norton & Neely, 1987). The click stimulus has a broad spectrum, and hence consequently it stimulates broad frequency region of the cochlea in a single measurement. Lutman, Mason, Sheppard and Gibbin (1989) opined that, the presence of click evoked OAEs (CEOAEs) are a powerful indicator of normal hearing. Therefore it has been applied as a general tool in universal neonatal hearing screening (UNHS) programs. Toneburst evoked OAEs (TBOAEs) include narrow bandwidth tone stimuli, which has stimulus energy concentrated on a particular area of the basilar membrane and elicits a more frequency-specific cochlear response. An advantage of toneburst stimulus is that more energy can be introduced in a specific frequency range compared to click, which is a more frequency dispersive stimuli. Although tone burst stimuli have greater frequency specificity compared to click stimuli, TBOAEs have not been routinely used in pediatric populations.

Most research works on TEOAEs have been performed using click stimulus. There have been many studies showing that the CEOAEs can be recorded in most normally hearing adults and in normal newborns and in babies admitted to a special care unit (Stevens, Webb, Hutchinson, Connel, Smith & Buffin, 1989). Uziel and Piron (1991) recorded CEOAEs in neonates ranging in age from a few days to 2 months after birth. The results showed that, in neonates, the emissions were stronger and covered a wider frequency range than those from the adults. Stevens, Webb, Hutchinson, Connel, Smith and Buffin (1990) recorded CEOAEs in 30 normal newborns and showed that 96% produced an emission at the highest level tested of 41 dB nHL. These results prove high degree of confidence that can be placed in the test when used to detect hearing impairments.

Probst, Coats, Martin and Lonsbury-Martin (1986) carried out a study on normal hearing young adults using spontaneous, click and toneburst evoked otoacoustic emissions. It was found that toneburst-evoked emissions were often more prominent than click evoked emissions and no spontaneous emissions were detected.

[1]Email: chandnimallik2009@gmail.com
[2]Lecturer in Audiology, Email: mamms_20@rediff.com

Highly similar peaks were present in the spectra of toneburst-evoked emissions within the range of toneburst spectra. Wit and Ritsma (1979) documented that the higher the frequency of the stimulus the smaller the response. TBOAEs at 1 kHz were more robust than CEOAEs in terms of emission response level and signal-to-noise ratio (SNR) at both 1 and 1.5 kHz frequency bands. The prevalence rate for CEOAE and TBOAE responses in these two frequency bands was significantly different (Zhang, McPherson, Shi, Tango & Wong, 2007).

Most of the research has been carried out using click stimuli, and also the studies have been concentrated on adult population. Very few studies have been carried out using toneburst stimuli for hearing screening and hearing assessment. As to the research and clinical application of TBOAEs, studies have been mainly focused on adult population. Whereas,there are very limited number of studies which have been done on infants. Hence there is a need to study the utility and efficacy of TBOAEs among infants. Also very limited studies have been done on the spectral information of TEOAEs as the age advances. The information regarding the comparison between CEOAEs and TBOAEs in infants is limited. Hence the present study was undertaken. The aims of the present study were, to investigate the influence of age on click and toneburst evoked OAEs; to investigate the pattern of frequency shifts with age; to study the sensitivity of click and toneburst evoked OAEs; and, to monitor the frequency specific maturational changes in the cochlea.

Method

Participants

The study was conducted on forty infants (80 ears) in the age range of 0 to 12 months. The forty infants were further categorized into four subgroups based on their age as, Group I: 0 to 4 months; Group II: 4 to 6 months; Group III: 6 to 8 months, and; Group IV: 8 to 12 months. Each age group had 10 infants each. All the infants had normal otoscopic examination indicating absence of external and middle ear pathology. They were healthy with no symptoms of cold or ear discharge at the time of assessment. They had no complaints and prior histories of any high risk factors and neurological symptoms. All the infants had age appropriate minimum response levels in behavioral observation audiometry, normal outer hair cell functioning ensured by recording TEOAEs and normal hearing sensitivity ensured by recording auditory brainstem responses (ABR).

Instrumentation

An otoscope was used to observe the status of external auditory canal and tympanic membrane. To present the stimuli for behavioral observation audiometry or visual reinforcement audiometry, a calibrated two-channel diagnostic audiometer Madsen Orbiter-922 (version 2) with impedance matched loudspeakers were used. A calibrated Grason Stadler Inc. Tympstar middle ear analyzer (GSI Tympstar, version 2) was used to carry out tympanometry and acoustic reflexometry. A personal computer based Intelligent Hearing Systems Smart EP version 3.94 evoked potential system to record ABR, and a calibrated otoacoustic emission system ILO version 6 software (Otodynamics Ltd., UK) was used to record TEOAEs for both clicks and tone burst stimuli.

Test Environment

All the tests were carried out in a sound treated room with noise levels within permissible limits as per ANSI S3.1 (1991). The test room was comfortable enough for the infants in terms of light and temperature.

Test Procedure

Case history: Detailed information regarding the history of prenatal, natal and postnatal medical conditions were collected for all the infants. A detailed report regarding the auditory behavior of infants at home for various environmental sounds like calling bell, voices from TV or radio, pressure cooker whistle etc was obtained from the parents or caregivers.

High Risk Register: Medical reports were reviewed to make sure that all the infants were devoid of various risk factors and other medical conditions. This was done by administering the modified high risk register developed by Anitha & Yathiraj (2001) to rule the high risk factors in infants.

Otoscopic Examination: The visual examination of the ear canal and the tympanic membrane of infant's ear were done using a hand held otoscope. This was done to rule out the presence of wax, foreign bodies in the ear canal and/or tympanic membrane pathologies.

Table 1: Parameters used to record click and toneburst evoked TEOAEs

Parameters	
Stimulus type	Click and Toneburst
Stimulus intensity	80 dBpk SPL
Stimulus rate	50 stimulus/sec
No. of stimulus	260
Stimulus polarity	Non-linear
Amplification	100-10,000 times
Filter settings	High pass filtered at 300 or 400 Hz

Tympanometry and Acoustic Reflex Measurements: Tympanograms were obtained using 1000 Hz probe tone frequency for infants till the age of 6 months and 226 Hz probe tone frequency for infants above 6 months of age. The pressure was swept from positive to negative (+200 to -400 daPa) with the pump speed of 200 daPa/sec and the probe intensity was 85 dB SPL for 226 Hz and 75dB SPL for 1000 Hz probe tone frequency. Ipsilateral acoustic reflex thresholds (ARTs) were measured for pure tone stimuli of 500 Hz, 2 kHz and 4 kHz using 226 Hz and 1 kHz probe tone frequencies. ART for 1 kHz was not recorded when 1 kHz probe tone was used as it might interact with the reflex activator signal frequency causing artifacts (Wilson & Margolis, 2001). ARTs were determined using an ascending technique by increasing the intensity of the activating stimuli in 5 dB steps from 60 dBHL as the starting intensity until reflex was obtained or equipment limit was reached. The minimum intensity at which the repeatable change in the admittance value is observed by taking the criterion as 0.03 mmhos was considered as an acoustic reflex threshold.

Behavioral Observation Audiometry: The behavioral responses of the infants were observed in the free field condition using warble tones from 500 Hz to 4000 Hz separated in octaves and for speech stimuli. The lowest levels of presentation of each of the stimuli, at which the subject exhibited some sort of auditory behavior was noted down.

Visual Reinforcement Audiometry: Thresholds for the infant's natural response (head turn) to the warble tones at octave frequencies from 250 Hz to 8 kHz and also for speech stimuli were obtained. Head turn towards the stimuli or the LCD TV monitor, was considered as a response only when it occurred within 3 seconds of the stimulus presentation.

Auditory brainstem response: ABR was recorded using IHS Smart EP system, version 3.94 with insert earphone (ER- 3A) using repetition rate of 11.1/s and filter setting of 30 to 3000 Hz with an analysis time of 15 ms. Initially, the electrode sites were cleaned with the help of skin preparing gel. Vertical electrode montage was used, where the non inverting electrode was placed on high forehead, inverting on the test ear mastoid and common on the non test ear mastoid. Electrodes were placed on the recording sites with the conduction paste and then were fixed with the help of surgical tape. It was ensured that the independent electrode impedance was less than 5 kΩ and inter electrode impedance was within 2 kΩ. The subjects were considered to have normal hearing sensitivity if the ABR wave V was clearly seen at 30 dB nHL

Transient evoked otoacoustic emissions (TEOAEs): Infants were tested during natural or sedated sleep or in quiet condition while recording TEOAEs. The probe was inserted gently into the ear canal with an appropriate sized probe tip so as to give a flat stimulus spectrum across the frequency range. Both click and tone burst-evoked OAEs were recorded. The evoked response for click stimulus included the frequency range from 500 Hz to 6 kHz, and the evoked response for toneburst was recorded individually at 500 Hz, 1 kHz, 2 kHz and 4 kHz. The stimulus was clear, with a positive and negative deflection followed by a flat line. The stimulus spectrum was smooth, with a rounded curve. The protocol used for recording click and tone burst evoked TEOAEs are mentioned in the Table 1.

The stimuli for clicks are trains of 4 biphasic clicks of 80 μs in the non-linear position. The non-linear protocol removes the stimulus artifacts of linear nature which can be misinterpreted as TEOAE response. In order to ensure a significant artifact rejection, the first 2.5 ms of the recording was eliminated. On termination of the test, the OAE response data that is obtained is averaged and an OAE waveform is displayed in the time domain. A Fast Fourier Transform (FFT) is performed on the OAE response spectrum and the OAE and noise energy are displayed in a frequency spectrum. The responses were considered as emissions based on the signal-to-noise ratio (SNR) and reproducibility. The overall SNR of greater than or equal to +3 dB and the reproducibility of greater than 50% were considered for the presence of otoacoustic emissions to determine normal outer hair cell functioning.

Results and Discussion

Absolute amplitude and SNR for different frequencies of CEOAEs and TBOAEs across age groups

The mean, SD and range values for absolute amplitude and SNR for both CEOAEs and TBOAEs at 1 kHz, 2 kHz and 4 kHz, across different age groups are depicted in the Figure 1 and 2. The results revealed that, the mean absolute amplitudes increase with increase in age for 1 kHz click frequency. But such a trend was not observed for 2 kHz and 4 kHz click frequencies. The mean absolute amplitude was more at 2 kHz and 4 kHz click frequencies compared to 1 kHz across all the age groups. Also, as it is evident from Figure 1, no definite trend was followed for the mean absolute amplitude across the age groups for different toneburst frequencies. However, the absolute amplitude was found to be higher at 2 kHz and 4 kHz compared to l kHz toneburst frequency (Figure 1). The mean SNR values also increased with increasing age only for 1 kHz click frequency. The mean SNR was more at 2 kHz and 4 kHz compared to 1 kHz across all the age groups (Figure 2). When SNR values for different toneburst stimulus frequencies were considered, no definite trend was observed. However, the mean SNR was found to be higher at 2 kHz and 4 kHz compared to l kHz toneburst frequency.

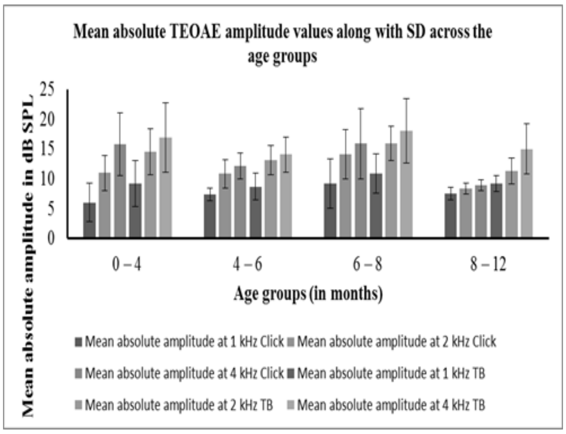


Figure 1: Mean, SD and range for absolute amplitude at 1 kHz, 2 kHz, and 4 kHz for click and tone burst stimuli across age groups.

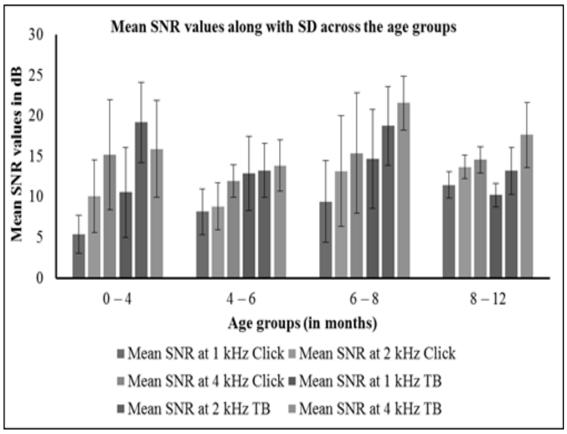


Figure 2: Mean, SD and range for SNR at 1 kHz, 2 kHz, and 4 kHz for click and tone burst stimuli across age groups.

A possible explanation for higher absolute amplitude at 2 kHz and 4 kHz is that, the resonance frequency of the infants ear canal is of high frequency. Thus at higher frequencies the absolute amplitude is more. As the age increases, the resonant frequency of the ear canal shifts towards the mid frequencies. This would have resulted in an increase in the absolute amplitude at 1 kHz with increase in the age. The reduction in the absolute amplitude at other frequencies with increasing age could be attributed to the fact that, with the increase in age, the ear canal volume increases and because the SPL measured inside the cavity is inversely proportional to the volume, this could have led to the decrease in absolute amplitude with increase in age.

The findings of the present study are in accordance with the findings stated by Smurzynski (1994). It was found that the click evoked TEOAE levels increases with the post conceptional age. The TEOAE levels were found to decrease with increasing age (Engdahl, Arnesen & Mair, 1994; Glattke, Pafitis, Cummiskey & Herer, 1995; Norton & Widen, 1990; Nozza & Sabo, 1992; Prieve, Fitzgenald, & Schuttle, 1997; Speak, Leonard, Kim, Jung & Smurzynski, 1991), which is also evident from the present study, where, TEOAE amplitudes were reduced for 8 to 12 months age group infants at all the frequencies. Kapoor and Panda (2006) analyzed SNR values for TEOAEs in neonates (0 to1 month) and infants (1 to 12 months). The results showed that the neonates had the lowest SNR ranging between 3.47 to 9.62 dB whereas the infants showed the highest SNR values ranging between 6.13 to 13.11 dB. These findings support the present study where similar results were found. The results of the present study can also be supported by results of the study done by Shi, Wang, Yuan and Zhang (2010), where they reported that lowest SNRs were obtained at low frequency (0.8 kHz) for neonates and at higher frequency (4 kHz) for younger adults.

Zhang et al. (2008) investigated the characteristics (amplitude and SNR) of the 1 kHz TBOAE response in neonates. The results showed that, the mean TBOAE response at 1 kHz was significantly higher than that of 2 kHz ($p < 0.05$), whereas, the SNR at 1 kHz was significantly lower than that at 1.5 kHz and 2 kHz frequency bands. The main reason for these findings relates to the greater noise in the 1 kHz frequency region, which may mask out the weaker TBOAE response. And also, a default low frequency filter setting in the ILO system may also reduce the noise levels at lower frequencies, and also reduces the OAE response. The mean TBOAE responses obtained at 1 kHz in the present study did not correlate with the results obtained from Zhang and McPherson (2008) study. Whereas, in the present study, the SNR values at 1 kHz were in accordance with Zhang and McPherson (2008) study.

Comparison of Absolute Amplitude and SNR across Age groups for Different Frequencies of CEOAEs and TBOAEs

Mixed ANOVA was carried out to see the significant interaction across age groups, frequencies and stimuli for amplitude and SNR measures separately. The results of mixed ANOVA for absolute amplitude revealed, a high statistically significant difference between the age groups [$F(3, 76) = 13.24$, $p < 0.05$]; between different frequencies [$F(2, 152) = 88.89$, $p < 0.05$]; between the two different stimuli [$F(1, 76) = 105.7$, $p < 0.05$] for all the four age groups; between frequencies and age groups [$F(6, 152) = 9.09$, $p < 0.05$]; between age groups and stimuli [$F(3, 76) = 1.53$, $p < 0.05$]; between stimuli and frequencies [$F(2, 152) = 8.37$, $p < 0.05$]; also between the stimuli, frequencies and age groups [$F(6, 152) = 3.52$, $p < 0.05$]. The results of mixed ANOVA for SNR revealed, a high statistically significant difference between the age groups [$F(3, 76) = 5.70$, $p < 0.05$]; between different frequencies [$F(2, 152) = 55.04$, $p < 0.05$]; between the two different stimuli [F

(1, 76) = 95.8, $p < 0.05$] for all the four age groups; between frequencies and age groups [F (6, 152) = 4.44, $p < 0.05$]; between age groups and stimuli [F (3, 76) = 4.35, $p < 0.05$]; between stimuli and frequencies [F (2, 152) = 10.04, $p < 0.05$]; and also between the stimuli, frequencies and age groups [F (6, 152) = 8.41, $p < 0.05$].

Effect of Age on Absolute Amplitude and SNR at Each Click Frequency

As mixed ANOVA showed significant interaction across click frequencies on amplitude and SNR measures, to see the effect of age on amplitude and SNR, multivariate analysis of variance (MANOVA) was done at each click frequency. The results for amplitude measures showed that there was significant difference in amplitude across age groups for 1000 Hz [F (3, 76) = 4.40, $p < 0.05$]; 2000 Hz [F (3, 76) = 13.61, $p < 0.05$]; and for 4000 Hz [F (3,76) = 13.39, $p <0.05$] click frequencies. The results for SNR measures showed that there was significant difference in SNR across age groups for 1000 Hz click frequency [F (3, 76) = 12.56, $p < 0.05$]; and for 2000 Hz click frequency [F (3, 76) = 5.89, $p < 0.05$]. There was no significant difference in SNR for 4000 Hz click frequency [F (3,76) = 1.86, $p > 0.05$] across age groups.

As MANOVA showed significant difference between age groups for 1000 Hz, 2000 Hz and 4000 Hz click frequency for amplitude measures, and significant difference between age groups for 1000 Hz and 2000 Hz click frequency for SNR measures, further analysis using the Duncan post hoc analysis test was done to see between which two age groups, the amplitude and SNR differ significantly. The results of Duncan post hoc test for absolute amplitude for each click frequency revealed that, for 1 kHz click frequency, group I was statistically significant from group III and vice versa. For 2 kHz click frequency, group III and IV significantly differed from all the age groups. However, group I and II significantly differed from both the groups III and IV. And for 4 kHz click frequency, group I was statistically significantly different from all the other age groups except group III. Similarly group II showed significant difference from all the age groups except from group IV. These results were similar to the results obtained by Norton et al. (2000) where, there was no much difference in the TEOAE levels for infants born between 28 and 40 weeks of GA and those between birth and 28 days after birth. And a significant group effect was observed at 2 kHz click frequency.

The results of Duncan post hoc test for SNR for each click frequency showed that, for 1 kHz click frequency, only group I was found to be significantly different from the other age groups. And group II was significantly different from group I and IV. Whereas, group III was significantly different from only group I. For 2 kHz click frequency, group I showed no statistically significant difference between the age groups. Whereas, group II was statistically significant from group III and IV. For 4 kHz click frequency, the SNR values did not show any significant difference across the age groups.

Effect of Age on Absolute Amplitude and SNR at each Toneburst Frequency

To see the effect of age on absolute amplitude and SNR measures, MANOVA was done at each tone burst frequencies. The results showed that, there was no significant difference in absolute amplitude [F (3, 76) = 2.41, $p > 0.05$] across age groups for 1000 Hz tone burst frequency. However significant difference in absolute amplitude across age groups for 2000 Hz and 4000 Hz tone burst frequency [F (3, 76) = 9.03, $p < 0.05$] and [F (3,76) = 2.94, $p < 0.05$] respectively was observed. A significant difference in SNR across age groups for 1000 Hz, 2000 Hz, and 4000 Hz tone burst frequency [F (3, 76) = 3.86, $p < 0.05$], [F (3, 76) = 12.86, $p < 0.05$], and [F (3,76) = 11.75, $p < 0.05$] respectively was also observed.

As MANOVA showed significant difference between age groups for 1000 Hz, 2000 Hz and 4000 Hz toneburst frequency on absolute amplitude and SNR, further analysis using the Duncan post hoc analysis test was done to see between which two age groups, the absolute amplitude and SNR differ significantly. The results of Duncan post hoc test for absolute amplitude for each toneburst frequency revealed that, 1 kHz toneburst frequency showed statistically no significant difference across the different age groups. For 2 kHz toneburst frequency, group I and II were statistically significant from group III only; whereas, group III was statistically significant from all the age groups, and group IV was significantly different from the other age groups except from group II. For 4 kHz toneburst frequency, group I and IV showed no statistically significant difference across the age groups. Whereas, group II was statistically significant from group III and vice versa.

The results of Duncan post hoc test for SNR for each toneburst frequency showed that, for 1 kHz toneburst frequency, group II showed no significant difference with the other age groups. When 2 kHz toneburst frequency was considered, the SNR values for group I was statistically significant from group II and group IV. And the group II showed statistically significant difference from all the other age groups except from group IV. Similarly, group III was statistically significant from all the other age groups except from group I. Group IV was statistically significant from all the other age groups except from group II. For 4 kHz frequency, group III was found to be significantly different from all the age groups. Group II and IV showed statistically significant difference from all the other age groups except from Group I. The findings of Zhang and McPherson (2008) were in support with the present study where, the age of the neonate (mean test age: 2.54 days) had no effect on the mean response and SNR at each frequency band (1 kHz, 1.5 kHz and 2 kHz). But this was not evident for all the frequencies in the present study.

Effect of click frequency on absolute amplitude and SNR at each age group

As mixed ANOVA showed significant interaction across click frequencies, repeated measure ANOVA was done at each age group to see the effect of age on amplitude and SNR. The results showed that there was significant difference in absolute amplitude across click frequencies in group I, II, III, IV [$F (2, 38) = 29.14$, $p < 0.05$], [$F (2, 38) = 46.49$, $p < 0.05$], [$F (2, 38) = 24.29$, $p < 0.05$], and [$F (2, 38) = 37.23$, $p < 0.05$] respectively. A significant difference in SNR values across click frequencies in group I, II, III, IV [$F (2, 38) = 26.61$, $p < 0.05$], [$F (2, 38) = 22.10$, $p < 0.05$], [$F (2, 38) = 8.17$, $p < 0.05$], and [$F (2, 38) = 20.02$, $p < 0.05$] respectively were also observed.

As repeated measure ANOVA showed significant difference between click frequencies across age groups on absolute amplitude and SNR, further analysis using the Bonferroni's multiple pair wise comparison was done to see between which two click frequencies, the absolute amplitude and SNR differed significantly. The results of Bonferroni's multiple pairwise comparison test for absolute amplitude, for each age group revealed that, for group I and IV, the absolute amplitude for the click stimulus of 1 kHz, 2 kHz and 4 kHz was significantly different from the other frequencies ($p < 0.05$). Whereas, group II showed statistically significant difference for 1 kHz and 2 kHz, but 4 kHz was significantly different only from 1 kHz, and for group III, only 1 kHz was significantly different from the other frequencies.The results of Bonferroni's multiple pair wise comparison test for SNR for each age group revealed that, group I was significantly different across all the frequencies. For group II, 1 kHz and 2 kHz showed significant difference only from 4 kHz click frequency. For group III and IV, the SNR values at 1 kHz were found to be statistically significantly from the other frequencies. Whereas, 2 kHz and 4 kHz was significantly different only from 1 kHz click frequency.

Effect of tone burst frequency on absolute amplitude and SNR at each age group

Repeated measure ANOVA was done at each age group to see the effect of age on absolute amplitude and SNR values. The results showed that there was significant difference in absolute amplitude across tone burst frequencies in group I, II, III, IV [$F (2, 38) = 21.15$, $p < 0.05$], [$F (2, 38) = 64.73$, $p < 0.05$], [$F (2, 38) = 24.61$, $p < 0.05$], and [$F (2, 38) = 52.04$, $p < 0.05$] respectively. The results of repeated measure ANOVA for SNR revealed, significant difference across click frequencies in group I, III, and IV [$F (2, 38) = 24.34$, $p < 0.05$], [$F (2, 38) = 14.61$, $p < 0.05$], and [$F (2, 38) = 58.33$, $p < 0.05$] were observed respectively. Whereas, group II showed no significant difference in SNR [$F (2, 38) = 1.39$, $p > 0.05$] across click frequencies.

As repeated measure ANOVA showed significant difference between tone burst frequencies across age groups on absolute amplitude and SNR, further analysis using the Bonferroni's multiple pair wise comparison was done to see between which two tone burst frequencies, the absolute amplitude and SNR differ significantly. The results of Bonferroni's multiple pair wise comparison test for absolute amplitude, for each age group revealed that, for group I, 1 kHz toneburst frequency was statistically significant from all the other frequencies and, 2 kHz toneburst frequency was not significantly different from 4 kHz frequency and vice versa. Whereas, statistically significant difference was observed between 2 kHz and 1 kHz, and 1 kHz and 4 kHz. For group II, the absolute amplitude at all the tone burst frequencies was statistically significant, except for the 4 kHz which did not have a significant difference with 2 kHz. Similarly, the 4 kHz tone burst frequency showed statistically no significant difference with 2 kHz frequency and vice versa for group III. There was statistically significant difference between 2 kHz and 1 kHz, and 4 kHz and 1 kHz ($p < 0.05$). Whereas, group IV showed statistically significant difference across all the tone burst frequencies. The results of Bonferroni's multiple pair wise comparison test for SNR, for each age group revealed that, all the age groups, except group II, showed a statistically significant difference ($p < 0.05$) across all the toneburst frequencies tested.

Comparison of absolute amplitude and SNR across click and tone burst stimuli at each age group for each frequency

As mixed ANOVA showed significant interaction across stimuli, paired sample t-test was done, in order to compare the absolute amplitude of click frequencies (1 kHz, 2 kHz & 4 kHz) and tone burst frequencies (1 kHz, 2 kHz & 4 kHz) at each age group for each of the frequencies. The findings for the comparison of absolute amplitude revealed that, for group I, there was statistically significant difference at 1 kHz [$t (19) = 2.69$, $p < 0.05$] and 2 kHz [$t (19) = 4.67$, $p < 0.05$] click and toneburst stimuli. Whereas, no significant difference was observed for the comparison of 4 kHz [$t (19) = 0.93$, $p > 0.05$] click and toneburst stimulus. For group III, statistically significant difference was observed only for 1 kHz click and toneburst frequencies [$t (19) = 2.43$, $p < 0.05$]. Statistically no significant difference was obtained for 2 kHz [$t (19) = 1.69$, $p > 0.05$] and 4 kHz [$t (19) = 1.49$, $p > 0.05$] click and toneburst stimuli. Whereas, for group II, statistically significant difference was observed at all the frequencies that is, at 1 kHz [t (19) = 3.54, $p < 0.05$], at 2 kHz [t (19) = 2.99, $p < 0.05$], and at 4 kHz [t (19) = 2.87, $p < 0.05$] click and toneburst stimuli. Similarly, for group IV, statistically significant difference was observed at all the frequencies, that is, at 1 kHz [t (19) = 4.46, $p < 0.05$], at 2 kHz [t (19) = 6.46, $p < 0.05$], and at 4 kHz [t (19) = 7.04, $p < 0.05$] click and toneburst stimuli.

The findings for the comparison of SNR revealed that, for group I, there was significant difference for the comparison of SNR values between click and toneburst stimuli at 1 kHz [t (19) = 3.96, $p < 0.05$] and 2 kHz [t (19) = 5.95, $p < 0.05$]. Whereas, 4 kHz click and toneburst stimulus showed no significant difference [t (19) = 0.34, $p > 0.05$]. Similarly, for group II, there was significant difference for the comparison of SNR values between click and toneburst stimuli at 1 kHz [t (19) = 3.94, $p < 0.05$] and 2 kHz [t (19) = 5.02, $p < 0.05$]. Whereas, 4 kHz click and toneburst stimulus showed no significant difference [t (19) = 1.98, $p > 0.05$]. For group III, the comparison of SNR values between click and toneburst stimuli was found to be significantly different at all the frequencies, that is, at 1 kHz [t (19) = 3.14, $p < 0.05$], at 2 kHz [t (19) = 4.87, $p < 0.05$], and at 4 kHz [t (19) = 3.20, $p < 0.05$] click and toneburst stimuli. Whereas, for group IV, except at 2 kHz [t (19) = 0.61, $p > 0.05$] click and toneburst stimulus, all the other frequencies, that is, 1 kHz [t (19) = 2.40, $p < 0.05$] and 4 kHz [t (19) = 3.69, $p <0.05$] click and toneburst stimuli showed a statistically significant difference.

Conclusions

From the present study it can be concluded that, TBOAEs showed a stronger response level than CEOAEs. Hence, to obtain a more frequency specific response, a toneburst stimulus can be used instead of click stimulus. A significant difference across the age groups was found for both CEOAEs and TBOAEs. The response in infants is more robust and it contains more high frequency energy. As the mean SNR value is more at higher frequencies for both click and toneburst stimulus, while evaluating the response in infants not only the amplitude of the response should be considered but also the SNR values.The values obtained from the present study can be used as a clinical tool for screening and diagnostic purposes in the pediatric population. Also, toneburst stimuli can be used instead of click stimuli to obtain a more frequency specific response in both pediatric and adult population. As the prevalence of TBOAEs are higher compared to CEOAEs in infants, it can be included in the screening and diagnostic test battery.

References

American National Standards Institute. (1991). *American National Standard maximum permissible ambient noise levels for Audiometric Test room.* (ANSI S3.1-1991). New York: American National Standards Institute.

Anitha & Yathiraj, A. (2001). *Modified high risk register (HRR) for professionals and non-professionals - formulation and its efficacy.* Unpublished masters dissertation, submitted to the university of mysore, mysore.

Engdhal, B., Arnesen, A. R., & Mair, I. W. S. (1994). Otoacoustic emission in the first year of life. *Scandinavian Audiology, 23,* 195-200.

Glattke, T. J., Pafitis, I. A., Cummiskey, C., & Herer, G. R. (1995). Identification of hearing loss in children and adults using measures of transient otoacoustic emission reproducibility. *American Journal of Audiology, 4,* 71-87.

Kapoor, R. & Panda, N. K. (2006). Transient evoked otoacoustic emissions. *Indian Journal of Pediatrics,73*(4), 283-286.

Kemp, D. T. (1978). Stimulated acoustic emissions from the human auditory system. *Journal of the Acoustical Society of America, 64,* 1386-1391.

Kemp, D. T. (1988). Developments in cochlear mechanics and techniques for non-invasive evaluation. *Advances in Audiology, 5,* 27-45.

Lutman, M. E., Mason, S. M., Sheppard, S., & Gibbin, K. P. (1989). Differential diagnostic potential of otoacoustic emissions: a case study. *Audiology, 28,* 205-210.

Norton, S. J., & Neely, S. T. (1987). Tone-burst evoked otoacoustic emissions from normal-hearing subjects. *Journal of the Acoustical Society of America, 81,* 1860-1872.

Norton, S. J., & Widen, J. E. (1990). Evoked otoacoustic emissions in normal-hearing infants and children: Emerging data and issues. *Ear and Hearing, 11,* 121-127.

Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., & Cone-Wesson, B. (2000). Identification of neonatal hearing impairment: evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear and Hearing, 21,* 508-528.

Nozza, R. J, & Sabo D. L. (1992). Transiently evoked OAE for screening school-aged children. *The Hearing Journal,* 29-31.

Prieve, B. A., Fitzgenald, T. S., & Schuttle, L. E. (1997). Basic characteristics of click-evoked otoacoustic emission in infants and children. *Journal of the Acoustical Society of America, 102,* 2860-2870.

Probst, R., Coats, A. C., Martin, G. K., & Lonsbury-Martin, B. L. (1986). Spontaneous, click and toneburst-evoked otoacoustic emissions from normal ears. *Hearing Research, 21,* 261-275.

Shi, J. F., Wang, N. Y., Yuan, J. J., Li, L., & Zhang, J. (2010). Comparison of transient evoked otoacoustic emissions in new born and frequency specific approach. *Zhonghua Er Bi Yan Hou. Jing wai ke za zhi, 45*(3), 206-211.

Smurzynski. J. (1994). Longitudinal measurements of distortion product and click evoked otoacoustic emissions of pre-term infants: Preliminary results. *Ear and Hearing, 15,* 210-223.

Speak, Z., Leonard, J., Kim, D. O., Jung, M. D. & Smurzynski, J. (1991). Otoacoustic emission in normal and Hearing Impairment Children and Normal Adults. *Laryngoscope, 101,* 965-976.

Stevens, J. C., Webb, H. D., Hutchinson, J., Connel, J., Smith, M. F., & Buffin, J. T. (1990). Click Evoked otoacoustic emissions in neonatal screening. *Ear and Hearing, 11,* 128-133.

Stevens, J. C., Webb, H. D., Hutchinson, J., Connel, J., Smith, M. F.,& Buffin, J. T. (1989). Click evoked otoacoustic emissions compared with brainstem electric response.*Archives of Diseases in Childhood, 64,* 1105-1111.

Uziel, A., & Piron, J. P. (1991). Evoked otoacoustic emission from normal newborn and babies admitted to an intensive care baby unit. *Acta otolaryngology, (supplement 482),* 85-91.

Wilson, R. H., & Margolis, R. H. (2001). *Reflexo Acoustica*. In F. E. Musiek, W. F. Rintelmann, (Eds.) Perseptivas atuais em avaliacao auditva (1st ed.)(pp. 127-161). Sao Paulo: Ed, manhde.

Wit, H. P., & Ritsma, R. J. (1979), stimulated acoustic emissions from the human ear. *Journal of the Acoustical Society of America,66,* 911-913.

Zhang, V. W., & McPhersan, B. (2008). Tone burst evoked otoacoustic emissions in neonates; normative data. *BMC ear, nose and throat disorders, 8,* 171-176.

Zhang, V. W. McPherson, B., Shi, B. X., Tango, J. L. F., & Wong, B. Y. K. (2007). Neonatal hearing screening: A combined click evoked and tone burst otoacoustic emissions approach. *International journal of Paediatric Otorhinolaryngology, 72,* 351-360.

Influence of Auditory Closure and Working Memory on Audio-Visual Perception of Speech

[1]Deepashree & [2]Sandeep Maruthy

Abstract

The present study investigated the influence of the speech perception in noise and the working memory on audio-visual integration in speech perception. Sixty six adult native speakers of Kannada in the age range of 18 to 70 years, having normal hearing and normal or corrected vision of 6/6, participated in the study. The participants were assessed for their speech perception in noise, working memory capacity and audio-visual integration. In the assessment of audio-visual integration there were three conditions in quiet (A, V, AV) and eight conditions in the presence of noise (A, AV in +5dB, 0dB, -5dB & -10 dB SNR). Results showed that both SPIN and working memory have positive effect on speech perception through AV-mode. The influence was particularly significant at lower SNRs and the relationship was direct. That is, AV speech perception was better in individuals who had better SPIN and/or working memory. Thus, it can be concluded that internal redundancy controlled by central auditory mechanism/s (as evident through SPIN) and the cognitive abilities (represented as working memory, a cognitive factor) has a direct relationship with the AV speech perception and hence demands attention while interpreting on AV speech perception.

Keywords:*Speech perception, working memory, auditory enhancement, visual enhancement.*

Introduction

Speech perception is influenced by visual cues even though it is primarily an auditory process. In instances where auditory cues are compromised (such as in noisy environments or hearing impairment), visual input has been reported to significantly improve speech intelligibility by supplementing the missing auditory cues (Tye-Murray, Sommers & Spehar, 2007; Munhall, Kroos, Jozan & Vatikiotis-Bateson, 2004; MacLeod & Summerfield, 1987). The benefits provided by bimodal presentation of a speech signal are larger when auditory stimuli are degraded than when the speech signal is clear (Sumby & Pollack, 1954; O'Neil, 1954; Neely, 1956; Erber, 1969; Grant & Seitz, 2000; Rudmann, Mc Carley & Kramer, 2003; Bernstein, Auer & Takayanagi, 2004; Ross, Saint-Amour, Leavitt, Javitt & Foxe, 2007). But studies on relationship between the amount of redundancy and audio-visual (AV) speech perception show equivocal results (Sumby & Pollack, 1954; Erber, 1969; Summerfield, 1979; Middelweerd & Plomp, 1984; Shannon, Zeng & Wygonski, 1998; Anderson, 2006; Huffman, 2007).

Sumby and Pollack (1954) demonstrated that with increase in difficulty of auditory-only perception, the benefits obtained by combining the auditory and visual speech information also increased. Erber (1969) reported an improvement of 60% in word recognition scores from auditory-only condition to AV condition at -10 dB SNR for young adults.Similar results were reported using sentence materials (Middelweerd & Plomp, 1984; Summerfield, 1979). Whereas Anderson (2006) and Huffman (2007) reported that the amount of AV integration did not vary across different auditory signal manipulations and hence, systematically removing information from the auditory stimulus does not necessarily affect the degree of integration benefit.

The studies on effects of age are equivocal. The study by Tye-Murray, Sommers and Spehar (2007) compared the dependency on visual cues between older adults with and without hearing loss and suggested that older adults with hearing loss may rely much more on the visual cues than those with normal hearing. The degree of hearing loss was also shown to affect the integration of auditory and visual syllables as reported by Grant, Walden and Seitz (1998). Further, the older adults are reported to be less successful in combining information across two or more sensory modalities (Shoop & Binnie,1979; Middelweerd & Plomp, 1984; Plude & Hoyer, 1985). But in contrast, Spehar, Tye-Murray and Sommers (2008) suggested that the intermodal and intra-modal integration abilities are largely resistant to changes with age and hearing loss. The researchers have argued that age related changes in cognitive or central auditory processing abilities play a limited role in the poor recognition of speech (Sommers, 1997; Schneider, Danema, & Pichora-Fuller, 2002).

It is shown in one of the earlier studies that older adults have decreased processing resources available to them and hence perform poorer than young adults on memory tasks (Craik & Byrd, 1982). It is suggested that there is an interdependence of aging sensory systems and cognitive functions (Li & Lindenberger, 2002). The construct of working memory is considered important in the study of cognitive aging even though there are other measures of processing resources such as perceptual speed and reasoning. Studies by Park and colleagues

[1]Email: nimmadeepa@gmail.com,
[2]Reader in Audiology, Email: msandeepa@gmail.com

(1996, 2002) measured the working memory capacity in adults across life span and it was found that it was greatest in young adulthood and then decreases across life span. Martin and James (2005) proposed that age-related deficits in processing of inter-hemispheric information may bring about some of the listening problems in older adults. Hence, it is suspected that age-related changes in working memory may provide basis for decreased age-related performance on a range of cognitive tasks.

It has been explained that speech understanding needs a constant encoding of information into and out of the working memory and manipulation of the information stored in memory (Pichora-Fuller, Schreider & Daneman, 1995). Hence, loss of working memory hinders speech communication and may contribute to the age-related declines in the comprehension of speech which may be evident even in favorable listening conditions. In adverse listening conditions such as in the presence of noise, this may degrade further (Cohen, 1979, 1981; Light et al., 1982; Pichora-Fuller, Schreider & Daneman,1995). On the contrary, some of the earlier investigators have argued that age related changes in cognitive or central auditory processing abilities play a limited role in the poor recognition of speech (Sommers, 1997; Schneider, Danema, & Pichora-Fuller, 2002).

Thus, it is clear that the dependency on the visual information increase with degradation of the auditory signal, even though there are mixed opinions about the effect of the amount of redundancy on integration. This means that subjects depend on the information in other modalities when external redundancy of the auditory modality is cut down. For participants with hearing loss and processing disorder, there are equivocal studies about the effect of internal redundancy on audio-visual speech perception. The studies on aging effects have been inconclusive due to poor control of the variables like hearing loss and visual problem. Further, there was no systematic study to understand the relationship between working memory and AV speech perception. Hence, the present study was taken up to scientifically study the relationship among working memory, speech perception in noise and AV speech perception.

Method

In the present study, quasi-experimental research design was used to test the null hypothesis that there is no effect of speech perception in noise and working memory on audio-visual speech perception. The following method was used to test the hypothesis.

Participants

Sixty six normal hearing adults in the age range of 18 to 70 years participated in the study. The selected participants had pure tone thresholds within 15 dBHL at octave frequencies between 250 Hz and 8 kHz (ANSI, 1996), and normal or corrected vision of 6/6. All the selected participants were native speakers of Kannada.

All the participants were assessed for their speech perception in noise and working memory (details available later in this chapter). Based on their performance in speech in noise and working memory tests, they were divided into 4 groups. Confidence intervals (95%) of speech in noise scores and the working memory scores of the 66 participants were used as cut-offs. The scores less than lower boundary were grouped as 'poor' and the scores higher than the upper boundary were grouped as 'good'. As a result, 4 groups were formed.

Group I, named as Low speech in noise (LowSPIN) group had 20 participants with poor speech perception in noise i.e., speech identification scores of less than or equal to 72% (lower boundary score of confidence interval) at 0 dB SNR.

Group II, named as High speech in noise (HighSPIN) group had 36 participants with good speech perception in noise i.e., speech identification scores of more than 80% (upper boundary score of confidence interval) at 0 dB SNR.

Group III, named as Low working memory (LowWM) group had 14 participants with poor working memory i.e., working memory scores of less than or equal to 69% (lower boundary score of confidence interval).

Group IV, named as High working memory (HighWM) group had 21 participants with good working memory i.e., working memory scores of more than 76% (upper boundary score of confidence interval).

A written consent was obtained from each participant prior to their inclusion in the study.

Test Stimulus

Six lists of bi-syllabic Kannada (Dravidian language spoken in the state of Karnataka, India) words having ten words in each list were developed specifically for the purpose of the present study.

To begin with, 300 bi-syllabic Kannada words, frequently spoken by native speakers were collected from recorded speech samples, news papers and media interviews. From this list, 23 words with clusters were omitted leaving 277 words in the list. These 277 words were then given to 15 native speakers of Kannada to rate them according to familiarity on a 3-point scale (unfamiliar- if they were unaware of the word, familiar, & very familiar- if the word occurred very frequently in conversation). Out of these, 169 words, which were rated 'very familiar' by more than 12 participants (80%) were considered for the next level.

The selected 169 words were audio recorded using a computer with adobe audition (version 1.5) software at a sampling frequency of 44,100 Hz and 16 bit digitization using a unidirectional microphone, in an acoustically treated room. An adult female native speaker of Kannada, who was a professional voice user, uttered the words. The stimuli were further edited for removal of noise and, hiss reduction and a gap of three seconds was introduced between consecutive words, using the same software. Root Mean Square (RMS) normalization was done for all words in order to minimize differences in the stimulus amplitude.

The variability with respect to audibility was reduced and in turn the homogeneity across spondaic words was increased using the standard procedure (Hirsh, Reynolds, and Joseph, 1954). Ten speech and hearing professionals who had Kannada as their native language and with a minimum of three years training were selected for this procedure. To begin with, speech recognition threshold (SRT) was found out for all ten subjects using the procedure by Tillman and Olsen (1973), a descending method for SRT measurement. After obtaining SRT, the paired-words were presented at +4, +2, 0, -2, -4 and -6 dBSL (ref: SRT), and the participants were asked to orally repeat the words. The whole list of 169 words was presented at each sensation level (SL) and it was randomized during each presentation. The responses obtained from the ten listeners were analyzed to identify 'Easy' and 'Hard' words. The words which were missed once or less by all listeners when the list was presented at +4, +2, 0, -2, -4 and -6 dBSL were considered as 'Easy words'. Whereas the words which were missed five or more times by all listeners when presented at +4, +2, 0, -2, -4 and -6 dBSL were considered as 'Hard words' (based on procedure by Hirsh et al., 1952). All the 'Easy' and 'Hard' words were eliminated to increase the homogeneity with respect to audibility. This finally resulted in sixty bisyllabic Kannada words.

The frequency of occurrence of all the speech sounds in the list of sixty bisyllabic Kannada words were calculated in order to find out whether the list of bisyllabic words was phonetically balanced. Although this was not mandatory for testing the objectives of the current study, the investigator felt that a phonetically balanced list would have been an advantage while generalizing the results of the present study. Hence, the frequency of occurrence of sounds in the list was tested using the method used by Ramakrishna, Nair, Chiplunkar, Ramachandran and Subramanian (1962). The probability of any speech sound was obtained by taking ratio of number of occurrence of speech sound in the list to the total number of speech sounds in it. The relative frequencies of speech sounds obtained in the present study were compared with relative frequencies of speech sounds obtained by Ramakrishna et al., 1962. Both the relative frequencies were similar and hence the present list was phonetically balanced. Finally, from the list of sixty words, six lists having ten words in each list were obtained. The total number of vowels and consonants in the list which were obtained from the Ramakrishna et al.'s method were distributed equally across the six lists. Thus, it was attempted to keep similar frequency of occurrence of vowels and consonants in all the six lists of bisyllabic words.

Audio and Video Recording

Five adult, female, native speakers of Kannada, who were speech and hearing professionals with clear articulation, were selected and the video of final six lists were recorded by a professional videographer using a digital video camera. A white screen was used as a background and the participants were instructed to avoid bright clothing to avoid distractions in the visual stimuli. The participants were also instructed to produce the words clearly without exaggerating the articulators, reduce eye blinks and avoid head movements, while recording. The words which were articulated unclearly were recorded twice.

Simultaneously, along with the video recording, the audio of the speech stimuli was recorded digitally for all the five participants. The audio recording was done using a collar microphone, at a sampling frequency of 44,100 Hz and 16 bit digitization, using Praat Software (version 5.1.31). Out of the 5 samples (recordings of 5 individuals) of stimuli, best sample was selected based on clarity of stimuli visually as well as auditorily. The audio and video recordings were done in a sound treated room.

The auditory stimuli were edited using adobe audition (version 1.5) software for noise and hiss reduction and a gap of five seconds was introduced between the words. All the words were normalized to a constant scaling factor. To test in the degraded conditions, all the six lists were further superimposed with speech noise at +5 dB, 0 dB, -5 dB and -10 dB SNRs.

The visual stimuli of the selected subject were edited using Windows movie maker software, to introduce a gap of five seconds between the words. The original audio of the video sample was muted and the auditory stimuli which was recorded using Pratt software was overlapped with the visual stimuli. This was necessary as the original audio was distorted and had high background noise.

To test for homogeneity in auditory and visual stimuli among the lists, all the lists were presented to five Kannada native speakers in three modalities; only Auditory (A), only Visual (V) and Auditory-visual (AV). The pilot comparison showed that the scores obtained for all the lists were similar.

Instrumentation and Test Environment

In the present study, a calibrated Madsen Orbiter-922 type I diagnostic audiometer with TDH-39 headphones was used to estimate the air-conduction thresholds and administer speech audiometry. A laptop computer with windows movie maker was used for video editing. The Pratt and Adobe Audition softwares were used for recording, editing and presenting the test stimuli. A digital video camera was used to record visual stimuli and a Sony MX78 omni-directional collar microphone was used to record auditory stimuli. Headphones were used to present auditory stimuli in 'A' mode and 'AV' mode.

All tests were administered in an acoustically treated room with noise levels at permissible limits (ANSI S3.1, 1991).

Test Procedure

The procedure started with preliminary evaluations which included case history, puretone audiometry and speech audiometry. For all the participants, the puretone air conduction thresholds (0.25, 0.5, 1, 2, 4, and 8 kHz) and speech recognition thresholds were obtained monaurally for both the ears. Only the individuals who fulfilled all the subject selection criteria were chosen for the study. After preliminary evaluation, the procedure included assessment of speech perception in noise, working memory and audio-visual integration.

Speech perception in noise (SPIN) was binaurally tested at 0 dB signal to noise ratio (SNR). The phonemically balanced (PB) word list developed by Yathiraj and Vijayalakshmi (2005) was used as signal and was presented along with the speech noise. The participants were asked to repeat the words and total number of words repeated correctly was noted down. The percentage of correct responses was calculated by dividing number of correctly repeated words by the total number of words and multiplying this factor by 100.

The procedure used to measure working memory capacity was adapted from versions of the operation span task used by Kane, Hambrick, Tuholski, Wilhelm, Payne and Engle (2004). Guidelines recommended by Conway, Kane, Bunting, Wilhelm and Engle (2005) were followed during administration and scoring. The operation span task consisted of 5 items and 20 elements. The number of elements per item varied from 2 to 6. Each element consisted of a mathematical operation which included addition and division, and a Kannada word (E.g. (6/3) + 7 = 8? 'Ka:ge'). The participant's task was to read the mathematical problem aloud, then say 'yes' or 'no' to indicate whether the given answer was correct or wrong, and then say the word. After all the elements in an item were presented, the participants were required to repeat all the words in that item. The difficulty of the items was randomized such that the number of elements was unpredictable at the outset of an item. The scoring for mathematical problem and words were done separately. The participants who scored less than seventeen out of twenty (80%) in mathematical problem were not considered for analysis as results of those subjects are not valid. For each correct item, one mark was given, only if all the elements were repeated correctly. If some of the elements were incorrect, the number of correct elements was divided by total number of elements. Finally, the scores of all five items were added and divided by five to obtain the final score for working memory. The scores for working memory ranged between 0 and 1.

The testing for the assessment of audio-visual integration was conducted in quiet as well as in the presence of noise. In quiet, stimuli were presented in 3 modalities (auditory only, visual only & auditory-visual). In the presence of noise, the stimuli were presented in 2 modalities (auditory only, & auditory-visual) at four different signal to noise ratios (+5dB, 0dB, -5dB and -10 dB SNR). The audio stimuli were presented binaurally through headphones at most comfortable level (MCL). The visual stimuli were presented from wide screen laptop with fifteen inches and the distance between the participant and laptop was maintained at fifteen inches. The audio of the video stimuli was muted while testing in V mode. Finally there were three conditions in quiet (A, V, AV) and eight conditions in the presence of noise (A, AV in +5dB, 0dB, -5dB & -10 dB SNR). In each condition, only one list of 10 bi-syllabic words out of six lists was presented and the lists were randomized when they were needed to be presented for the second time. The mode of presentation was randomly chosen. The subjects responded by repeating the words and the words repeated were noted down by the clinician.

Response Analyses

In each condition (quiet, -10 dB, -5 dB, 0 dB, & +5 dB SNR) the total number of words repeated correctly out of ten words was noted separately for A, V, and AV mode. The AV speech perception was quantified by calculating visual enhancement (VE, the benefit obtained from adding a visual signal to an auditory stimulus) and auditory enhancement (AE, the benefit obtained from adding an auditory signal to a visual-only stimulus) scores. The VE and AE were calculated according to the equations 1 and 2 respectively. The scores for VE and AE ranged between -1 to +1.

$$VE = \frac{AV - A}{1 - A} \tag{1}$$

$$AE = \frac{AV - V}{1 - V} \tag{2}$$

The data obtained from the participants were subjected to the statistical analysis to test the objectives of the study.

Results

The primary objective of the study was to analyze whether the SPIN and WM influence AV perception of speech. The independent variables were SPIN and WM, while the dependent variable was AV speech perception. SPSS (version 16) was used for the statistical analysis of data. Descriptive statistics, Repeated measures ANOVA, correlation, independent t-test and one-way ANOVA were the statistical tests used for the purpose.

Effect of Modality on Speech Perception Scores

To begin with, the speech perception scores of 66 participants in the three modalities were analyzed to obtain the mean and standard deviation scores. This was done separately for each stimulus condition (-10 dB, -5 dB, 0 dB, +5 dB SNR, & quiet). The mean (raw score out of 10) and standard deviation of speech perception scores in the 3 modalities at different signal-to-noise ratios are given in Table 1.

In general, the mean speech perception scores were higher in AV-mode compared to A and V-mode. The scores were least in the V-mode. This was true in quiet as well as in all the SNR conditions. The mean scores also decreased with decrease in SNR in the A and AV modality. In the visual modality, only the scores obtained in quiet was used for modality-wise comparison in all the stimulus conditions.

The difference in the mean scores was tested for the statistical significance on Repeated measures ANOVA. The results of Repeated measures ANOVA are given in Table 2

Table 1: Mean and standard deviation (S.D) speech perception scores in the visual, auditory, and auditory-visual modes in quiet, -10 dB, -5 dB, 0 dB, 5 dB SNR

Stimulus Condition	V-mode Mean (S.D)	A-mode Mean (S.D)	AV-mode Mean (S.D)
-10 dB SNR	CNT	1.12 (0.87)	3.65 (1.96)
-5 dB SNR	CNT	3.58 (1.61)	6.82 (1.66)
0 dB SNR	CNT	7.74 (1.25)	9.21 (0.94)
5 dB SNR	CNT	9.71 (0.70)	9.85 (0.40)
Quiet	1.09 (1.17)	9.95 (0.21)	10.0 (0.00)

Note: CNT- could not be tested as presentation of auditory stimulus was inevitable in these conditions.

Table 2: Results of repeated measures ANOVA comparing speech perception scores across the 3 modalities in the 5 stimulus conditions

Stimulus Condition	F	df(error)
-10 dB SNR	108.881*	2(130)
-5 dB SNR	363.733*	2(130)
0 dB SNR	1718.676*	2(130)
5 dB SNR	3216.879*	2(130)
Quiet	3762.289*	2(130)

*Note: * - $p < 0.01$*

The results of repeated measures ANOVA showed a significant ($p<0.01$) main effect of modality on speech perception scores in all the 5 stimulus conditions. As there was a significant main effect, the pair-wise comparison was tested on Bonniferroni post-hoc test. The results of post-hoc test showed significant difference across all three modalities (A-V, A-AV & V-AV) at 0, -5 and -10 dB SNR. However, in quiet and +5 dB SNR the significant difference was found only between A-V and V-AV modalities.

Effect of Stimulus Condition on AV Speech Perceptions

In order to test the effects of independent variables on AV speech perception, the scores were first converted as ***visual enhancement*** and ***auditory enhancement*** scores. The visual enhancement VE score was a quantity of the benefit obtained from adding a visual signal to an auditory stimulus score, whereas the auditory enhancement AE score was a quantity of the benefit obtained from adding an auditory signal to a visual-only stimulus. The raw scores obtained in the 3 modalities were used to compute VE and AE scores, separately for each individual. This was done for all the 5 stimulus conditions. The mean and standard deviation of VE and AE scores in the 5 stimulus conditions are shown in Figure 1.

The inspection of the Figure 1 shows that there are mean differences among the 5 conditions in both AE and VE scores. The mean VE scores were maximum at 0 dB SNR and decreased with either increase or decrease in SNR. The AE scores on the other hand increased with increase in SNR. The mean differences observed were statistically tested for the effect of condition on repeated measures ANOVA. The results showed a significant main effect of condition on VE [$F(4,260) = 46.907$, $p<0.001$] and AE scores [$F(4,260) = 491.085$, $p<0.001$]. Consequently, pair-wise comparison was tested using Bonferroni post-hoc test and the results showed a significant difference in the mean scores across all conditions in AE. However in VE, there was no significant difference between quiet and +5 dB SNR, and, 0 and +5 dB SNR.

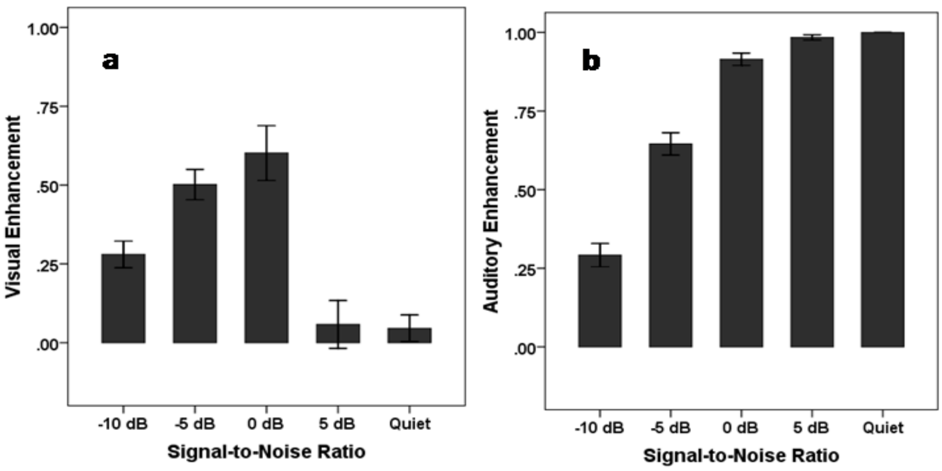


Figure 1: Mean and standard deviation of visual enhancement (a) and auditory enhancement (b) scores in quiet, -10 dB, -5 dB, 0 dB, and +5 dB SNR.

Relation between Speech Perception in Noise and AV Speech Perception

To understand the relation between SPIN and AV speech perception, the correlation was obtained for both VE and AE with SPIN scores in different stimulus conditions. The data were statistically tested for correlation between two variables on Pearson product moment correlation and the results are given in Table 3 and Table 4. The results of the correlation showed a significant low positive correlation at -10 dB and -5 dB SNR but a significant low negative correlation in quiet condition between mean scores of SPIN and VE. On the contrary, correlation of mean differences between SPIN and AE showed that there was a significant low positive correlation at -10 dB SNR but a significant low negative correlation in quiet and 5 dB SNR conditions. There was no significant correlation in other conditions for both VE and AE with SPIN scores.

The influence of SPIN on AV speech perception was further tested using a second method wherein the participants were divided into two groups based on their SPIN scores (good SPIN & poor SPIN groups) and compared with each other for their AV speech perception. The grouping was done based on the confidence interval (at 95%) determined for the SPIN scores of 66 participants of the study. The lower-boundary and the upper-boundary of the interval for the SPIN scores thus obtained were found to be 72% and 82% respectively. The participants with scores lesser than the lower boundary (72%) were grouped as 'poor SPIN' and those with scores higher than the upper boundary (82%) were grouped as 'good SPIN'. The participants with scores within the interval (72-82%) were not considered for any comparisons.

The VE and AE scores of the resultant two groups were then compared to investigate whether there was any difference in the AV speech perception between the two groups. It was hypothesized that the presence of significant difference in VE scores would indicate that SPIN influences AV speech perception. The mean and the standard deviation of the VE and AE scores of the two SPIN groups (good & poor), in each stimulus condition (quiet, -10 dB, -5 dB, 0dB & 5 dB SNR) are shown in Figure 2.

The mean VE scores were higher in the good SPIN group compared to poor SPIN group except in +5 dB SNR and quiet, wherein the case was reverse. In contrast, the mean AE scores of good SPIN group was

Table 3: The results of correlation test correlating visual enhancement (VE) with SPIN in quiet, -10 dB, -5 dB, 0dB, and +5 dB SNR

Stimulus Condition	SPIN and VE	
	Pearson Correlation	Significance
-10 dB SNR	0.496**	0.000
-5 dB SNR	0.253*	0.041
0 dB SNR	0.003	0.979
5 dB SNR	-0.166	0.184
Quiet	- 0.407**	0.001

*Note: * - $p < 0.01$*

Table 4: The results of correlation test correlating auditory enhancement (AE) with SPIN in quiet, -10 dB, -5 dB, 0dB, and +5 dB SNR

Stimulus Condition	SPIN and VE	
	Pearson Correlation	Significance
-10 dB SNR	0.512**	0.000
-5 dB SNR	0.335**	0.006
0 dB SNR	0.243*	0.049
5 dB SNR	0.207	0.095
Quiet	CNT	CNT

*Note: * - $p < 0.01$*

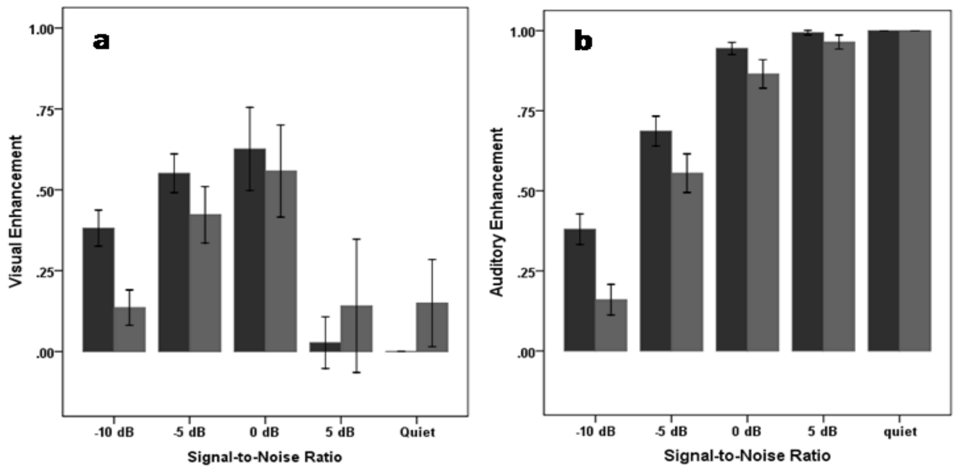


Figure 2: Mean and standard deviation of visual enhancement (a) and auditory enhancement (b) scores in quiet, -10 dB, -5 dB, 0 dB, and +5 dB SNR. Note: Visual and auditory enhancement scores would vary between -1 and 1 and WM scores ranged between 0 and 1.

Table 5: The results of the t-test comparing the good SPIN and poor SPIN groups for their visual enhancement (VE) and auditory enhancement (AE) scores in quiet, -10 dB, -5 dB, 0dB and +5 dB SNR

Stimulus Condition	VE(df= 54) t	AE (df= 54) t
-10 dB SNR	4.755**	4.967**
-5 dB SNR	2.042**	2.814**
0 dB SNR	0.557*	3.119**
5 dB SNR	-1.003	2.585*
Quiet	-2.475*	CNT

*Note: *- $p < 0.05$, **- $p < 0.01$, CNT- could not be tested as both the groups had same mean and no standard deviation.*

higher than that of poor SPIN group in all the stimulus conditions except in quiet, where the mean scores were equal. The mean scores of the two groups were statistically compared using independent samples t-test. The results of the t-test for VE and AE scores are given in Table 5.

The results showed that the scores for VE of good SPIN group were significantly higher in -10 dB and -5 dB SNR conditions and significantly lower in quiet, than the poor SPIN group. The mean differences in other stimulus conditions (0 dB & +5dB SNR) were not significant. On the contrary, the difference in the mean AE scores of the two groups was statistically significant in all stimulus conditions except in quiet

Relation between Working Memory (WM) and AV Speech Perception

The second independent variable of the study was working memory. Similar to effect of SPIN, the effect of working memory on AV speech perception was examined. The relation between WM and AV perception of speech across different stimulus conditions (quiet, -10 dB, -5 dB, 0dB, & +5 dB SNR) was examined by obtaining the correlation. The relationship between VE and AE with WM scores were statistically tested on Pearson product moment correlation test for different stimulus conditions (quiet, -10 dB, -5 dB, 0dB, and +5 dB SNR). The results of correlation test are given in Table 6 and Table 7.

From Table 6, we can witness clearly that, there was a significant moderate positive correlation at -10 dB SNR and a significant low negative correlation in quiet and 0 dB SNR conditions between WM and VE. Whereas, correlation between WM and AE showed a significant moderate positive correlation at -10 dB SNR and a significant low positive correlation at 0 dB SNR conditions as shown in Table 7.

Comparison between Poor and Good WM Groups

In a second method of testing the relationship between WM and AV speech perception, the participants were divided into good WM and poor WM groups to test whether there was any difference between the two groups. The confidence intervals (95%) for WM scores

Table 6: The results of correlation test comparing visual enhancement (VE) with working memory in quiet, -10 dB, -5 dB, 0dB, and +5 dB SNR

Stimulus Condition	WM and VE	
	Pearson Correlation	Significance
-10 dB SNR	0.572**	0.000
-5 dB SNR	0.192	0.178
0 dB SNR	-0.307*	0.028
5 dB SNR	-0.208	0.143
Quiet	- 0.349*	0.012

*Note: *- $p < 0.05$, **- $p < 0.01$, CNT- could not be tested as both the groups had same mean and zero standard deviation.*

Table 7: The results of correlation test comparing auditory enhancement (AE) with working memory in quiet, -10 dB, -5 dB, 0dB, and +5 dB SNR

Stimulus Condition	WM and AE	
	Pearson Correlation	Significance
-10 dB SNR	0.517**	0.000
-5 dB SNR	0.256	0.070
0 dB SNR	-0.330*	0.018
5 dB SNR	-0.040	0.781
Quiet	CNT	CNT

*Note: *- p < 0.05, **- p < 0.01, CNT- could not be tested as both the groups had same mean and zero standard deviation.*

were found out first. Then the participants with scores lesser than the lower boundary (70%) were grouped as 'poor WM' and those with scores higher than the upper boundary (76%) were grouped as 'good WM'. The participants with scores within the interval (70-76%) were excluded. The mean and standard deviation of VE and AE scores obtained for the 2 groups, across 5 conditions, are represented in Figure 3.

Figure 3 shows that the mean VE scores for the good WM group were better than those of poor WM at -10 dB, -5 dB and 0 dB SNR. But at +5 dB SNR and in quiet, poor WM group had better scores. Whereas, the mean AE scores of good WM group was higher than that of poor WM group in all the stimulus conditions except in quiet, where the mean scores were equal.

To test for significance in the mean differences between the two groups, independent samples t-test was administered. The results of the t-test for VE and AE scores across 5 conditions are given in Table 8. The results of the independent samples t-test showed that the scores for VE and AE of good WM group were significantly different than that of poor WM group only in -10 dB SNR condition. Even though there were mean differences in other conditions, they were not statistically significant.

Relation between Speech Perception in Noise and Working Memory

While analyzing the effect of WM on AV speech perception, across different SNR conditions, there is possibility of SPIN influencing the effect of WM. Hence, the relationship between the 2 independent variables was tested on correlation.

The Pearson product moment correlation was found out to test for correlation between SPIN and WM groups. The results showed that there was a significant ($p<0.001$) moderate positive correlation ($r=0.508$) between the two independent variables.

Table 8: The results of the t-test comparing the visual enhancement and auditory enhancement scores for good WM and poor WM groups in quiet, -10 dB, -5 dB, 0dB and +5 dB SNR

Condition	AE (df: 33)t	VE(df: 33)t
-10 dB SNR	2.380*	3.421**
-5 dB SNR	2.007	1.986
0 dB SNR	1.312	-0.986
+5 dB SNR	1.338	-0.337
Quiet	CNT	-1.817

*Note: *- p < 0.05, **- p< 0.01, CNT- could not be tested as both the groups had same mean and zero standard deviation.*

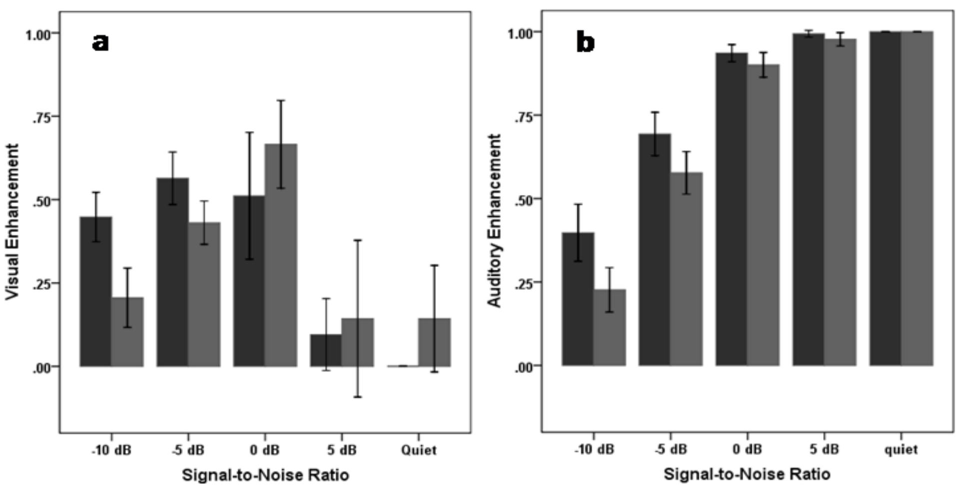


Figure 3: Mean, standard deviation of visual enhancement (Left panel) and auditory enhancement (Right panel) scores for good working memory (Dark bars) and poor working memory (Light bars) groups in the 5 different stimulus conditions. Note: Visual and auditory enhancement scores would vary between -1 and 1 and WM scores ranged between 0 and 1.

Discussion

The perception of speech could be through either auditory (A-mode) or visual (V-mode) mode. Of the two modalities, auditory modality is the primary one as it provides information on place, manner and voicing of speech sounds. The dependency on the visual mode is only in instances where the information provided through auditory modality is insufficient. This happens due to reduction either in external redundancy or internal redundancy. The negative influences of reduction in both external and internal redundancy on speech perception have been well established in the literature. Although it is universally accepted that audio-visual mode (AV-mode) is beneficial over A-mode in adverse listening conditions, the underlying mechanisms of AV perception are not clearly understood.

The present study was an attempt to probe into the underlying mechanisms of AV perception in sensory and cognitive domain. In the sensory domain auditory closure was taken as an independent variable, while in cognitive domain working memory was the independent variable. The present experiment to analyze the effects of these two independent variables on AV perception, on 66 participants, showed some interesting findings.

Effect of Modality on Speech Perception

Evidence from previous research explains that speech intelligibility improves when listeners receive information from both auditory and visual modes in instances where the auditory cues are degraded by reducing external redundancy (Sumby & Pollack, 1954; Anderson, 2006; Huffman, 2007). The present findings are in consonance with these earlier reports. Overall performance of the participants was best in the AV-mode compared to auditory-alone or visual-alone modes. The performance was least in the V-mode. Auditory mode provides cues pertaining to the place, manner and voicing of a speech sound while in the visual mode one gets cues of only place of articulation, that too not completely (Anderson, 2006). Hence, it is logical to expect better performance in the A-mode compared to V-mode. Further, as the redundancy of the cues is more in the bimodal presentation, AV-mode showed better performance than that in A-mode or V-mode, which were unimodal presentations.

The results showed that the performance in the A-mode was comparable to that in AV-mode in quiet and at +5 dB SNR. This means that the necessary cues for an ideal performance are not cut in quiet and +5 dB SNR. However, the performance reduced with further reduction in the SNR both in A and AV-modes. As the reduction of redundancy was only in the auditory signal without distorting the visual signal, the reduction in the scores of A and AV-modes can be attributed solely to the reduction in the SNR of the auditory signal. The Figure 4 represents the deterioration in the speech identification in A and AV-modes across different stimulus conditions.

As we can observe in the Figure 4, the performance of both A and AV- modes is same in quiet and decreases with decreasing SNR. But the decrease in performance for A-mode was more than that of AV-mode which is evident by the steepness of the curves. That is, the steepness of the A-mode (blue line) is greater than that of the AV-mode (red line). The difference between the A and AV-mode is due to complementary cues provided by the addition of visual cues, which however functions nonlinearly.

The increase in the difference between the performances in A and AV-modes with the decrease in SNR evidences, greater dependency on visual cues at lower SNRs. The finding is in agreement with the earlier studies (Sumby & Pollack, 1954; O'Neil, 1954; Neely, 1956; Erber, 1969; Grant & Seitz, 2000; Rudmann, Mc Carley & Kramer, 2003; Bernstein, Auer & Takayanagi, 2004; Ross, Saint-Amour, Leavitt, Javitt & Foxe, 2007) which showed larger benefits of bimodal presentations when the auditory signal was degraded. In the study by Sumby and Pollack (1954), it was demonstrated that the addition of visual cues improved speech perception by an amount equivalent to a 5 to 18 dB increase in the SNR which accounts for improvement of up to 60% in word recognition. They also demonstrated that with increase in difficulty of auditory-only perception, the benefits obtained by combining the auditory and visual speech information also increased. Although the present results showed a similar trend, the quantity of improvement at the worst SNR (-10 dB SNR) was only about 25%. The difference among the studies in the extent of improvement may be attributed to the different test materials used.

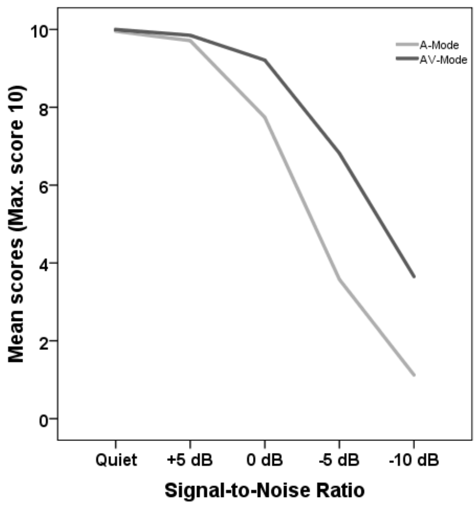


Figure 4: Mean identification scores in auditory mode (lower line) and audio-visual mode (upper line) in quiet, -10 dB, -5 dB, 0 dB, and +5 dB SNR conditions.

Effect of Stimulus Condition on AV Speech Perception

In the previous section, it was learnt that AV-mode was better than A-mode only at lower SNRs. That is, although the redundancy of visual cues was not varied, it nonlinearly facilitated speech perception. Hence the absolute difference between the scores of AV and A-modes would not have given a clear picture of the VE. For example, the importance of 10% facilitation by addition of V-mode will be different for 3 individuals who score 20%, 50% and 90% in the A-mode. An absolute difference between A and AV-mode would be 10% in all these cases, but logically the weightage of relative enhancement should be maximum for one who had 90% score in the A-mode followed by one with 50% and 20% score. This was achieved by calculating visual enhancement scores using the following formula in equation 3.

For the above example, the visual enhancement will be 0.12, 0.2, and 1 for the individuals with 10%, 50% and 90% score in the A-mode, respectively. With the similar logic the AE scores were derived from the performances in AV and V-mode using the equation 4. The AE score is a quantity of the benefit obtained from adding an auditory signal to a visual-only stimulus.

$$VE = \frac{AV - A}{1 - A} \tag{3}$$

$$AE = \frac{AV - V}{1 - V} \tag{4}$$

Effect of Stimulus Condition on Visual Enhancement (VE)

The results showed that the VE was maximum at 0 dB SNR and decreased with either increase or decrease in SNR. The decrease in the visual enhancement at SNRs below 0 dB was an unlikely finding considering the trend observed from the absolute scores obtained in AV and A-modes. The absolute difference between scores of A and AV-mode indicated increasing facilitation with decreasing SNR. However, the results of VE are in contrary to this. The trend of VE scores represents the functional benefit obtained by an individual by adding the visual cues. The functional benefit progressively reduces with reduction in the performance of A-mode, which in turn is dependent on the SNR.

On the other hand, the same logic does not apply to the performances at +5 dB SNR and quiet. Above 0 dB SNR, performance again reduces with increasing SNR. This is attributed to the ceiling effect. Because the scores in the A-mode were already 100%, the resultant VE scores would become 0. The number of such individuals would increase with increase in SNR, resulting in progressively decreasing mean VE performance. Based on these findings, it may be inferred that any study intending to investigate visual enhancement shall take SNRs of 0 dB or lesser and not take the higher SNRs, in order to get a true picture of VE.

Table 9: The Pearson correlation co-efficient correlating auditory enhancement scores and auditory mode scores in quiet, -10 dB, -5 dB, 0 dB, and +5 dB SNR conditions

Stimulus conditions	Correlation co-efficient
-10 dB SNR	0.352**
-5 dB SNR	0.423**
0 dB SNR	0.536**
5 dB SNR	0.488**
Quiet	CNT

*Note: **- p<0.001, CNT- could not be tested as all the participants had the same AE score of one (1).*

Effect of Stimulus Condition on Auditory Enhancement (AE)

The AE was maximum in quiet (approximately 40 dB SNR) and decreased with decreasing SNR. This trend is logical as the enhancement shall increase with increasing redundancy of the auditory cues. To strengthen the notion, the correlation between A scores and AE scores was done in different SNRs. It was expected to get high positive correlation between the two variables in all the stimulus conditions if the stated logic was true. Table 9 gives the Pearson correlation co-efficients in the 5 stimulus conditions.

The results showed that the correlation was maximum at 0 dB SNR and reduced at other conditions. Hence the above stated logic was defeated. The trend in Table 9 in turn shows the influence of VE as a primary variable. In the formula used for the calculation of AE, AV and V scores are the two parameters considered. As the V scores remained constant in all the stimulus conditions, the differences in the AE had to be because of AV which in turn was related to VE.

As the trend of VE (Figure 1) across conditions is same as that of trend of correlation observed in Table 9, it can be inferred that it is this influence of VE over AV scores that varied AE across stimulus conditions. This was supported by the results of correlation of VE scores and AE scores at 0, -5 and -10 dB SNR. The data of +5 dB SNR and quiet were not considered as VE in these conditions was erroneous due to ceiling effect. The results of correlation showed a moderate positive co-efficient ($r=0.632$, $p<0.001$) between AE and VE scores.

Effect of SPIN on AV Speech Perception

From the previous two sections, it is clear that AV speech perception is influenced by external redundancy of the auditory signal. Earlier studies have shown negative effects of advancing age on AV speech perception (Shoop & Binnie, 1979; Middelweerd & Plomp, 1987; CHABA, 1988; Walden, Busacco, & Montgomery, 1993; Cienkowski, 1999; Cienkowski & Carney, 2002; 2004). This age related decline could be either partially or completely due to reduced internal redundancy secondary to central auditory processing deficits. Hence it was of interest to study the relationship between internal redundancy and speech perception in noise. The relationship was tested using two methods; one, by correlating SPIN and AV speech perception. The other, by comparing the AV speech perception of good and poor SPIN groups.

The results of the correlation showed that below 0 dB SNR both VE and AE correlated with the SPIN scores. That is, VE and AE increases with increase in SPIN scores. Additionally, AE positively correlated with SPIN at 0 dB. A negative correlation between VE and SPIN is erroneous as most of the individuals had 0 VE at +5 dB SNR and quiet. Further, the comparison of AV speech perception among the good and poor SPIN groups also showed that the SPIN scores have a positive effect on speech perception through AV-mode.

From these findings it can be concluded that internal redundancy controlled by central auditory mechanisms, as evident through SPIN, is directly related to the AV speech perception. Hence, one should take SPIN into consideration while commenting on the benefit provided by AV speech perception.

Effect of Working Memory (WM) on AV Speech Perception

It has been explained that speech understanding needs a constant encoding of information into and out of the working memory and manipulation of the information stored in memory (Pichora-Fuller, Schreider & Daneman, 1995).

Similar to SPIN, the relationship between working memory and AV speech perception was analyzed by correlating the two parameters and by comparing good and poor working memory groups. This was done to decipher the underlying mechanisms of speech perception under the cognitive domain. Considering that the integration of information from two different modalities is a complex higher level task, cognitive factors such as working memory were expected to influence the AV speech perception.

The results of both correlation and group comparison showed evidence for working memory influencing AV speech perception in noise. The influence was particularly significant at lower SNRs and the relationship was direct. That is, AV speech perception was better in individuals who had better working memory.

The results of the present study are in consonance with the earlier report (Pichora-Fuller, Schreider & Daneman, 1995) which stated that the loss of working memory hinders speech communication and may contribute to the age-related declines in the comprehension of speech which may be evident even in favorable listening conditions.

From these findings it can be concluded that working memory, a cognitive factor, has a direct relationship with the AV speech perception and hence demands attention while interpreting on AV speech perception.

Relationship between SPIN and WM

As both working memory and speech perception in noise were influencing AV speech perception in a similar way, it was of interest to know how these two independent variables related to each other. The results of correlation showed that the speech perception in noise was better in individuals with good working memory and vice versa. However, it was only a moderate correlation which shows that they are not controlled by same physiological mechanism.

Conclusions

The present study has important implications in rehabilitative audiology. In audiological clinics, AV-mode of speech perception is recommended for individuals with poor speech identification scores. The present findings showed that the WM and SPIN directly influence AV speech perception. Hence, one should consider WM and SPIN, assess for them, estimate the probable benefit with AV speech perception and accordingly recommend the same.

References

American National Standards Institute. (1991). *American National Standard Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms*. ANSI S3.1- (1991). New York: American National Standards Institute.

Anderson, E. (2006). *Speech Perception with Degraded Auditory Cues.* Undergraduate honors thesis, The Ohio State University.

Bernstein, L. E., Auer, E. T., Jr., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*(1), 5-18.

Binnie, C. A., Jackson, P., & Montgomery, A. (1976). Visual intelligibility of consonants: A lipreading screening test with implications for aural rehabilitation. *Journal of Speech and Hearing Disorders, 41,* 530-539.

CHABA, Committee on Hearing and Bioacoustics, Working Group on Speech Understanding and Aging. (1988). Speech understanding and aging. *Journal of the Acoustical Society of America, 83,* 859?895.

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear and Hearing, 23,* 439?449.

Cienkowski, K. M., & Carney, A. E. (2004). The Integration of Auditory-Visual Information for Speech in Older Adults. *Journal of Speech-Language Pathology and Audiology, 28*(4), 169-172.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology, 11,* 412-429.

Cohen, G. (1981). Internal reasoning in old age. *Cognition, 9,* 59-72.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user?s guide. *Psychonomic Bulletin and Review, 12*(5), 769-786.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes*, (pp. 191-211), New York: Plenum.

Erber, N. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research, 12,* 423?425.

Grant, K. W., & Seitz, P.F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America, 4*(4), 2438-2449.

Grant, K. W., & Seitz, P. F. (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. *Journal of the Acoustical Society of America, 107,* 1000?1011.

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *Journal of the Acoustical Society of America, 103,* 2677?2690.

Hirsh, I. J., Reynolds, E. G., & Joseph, M. (1954). Intelligibility of different speech materials. *Journal of the Acoustical Society of America, 26,* 530?537.

Huffman, C. (2007). *The Role of Auditory Information in Audiovisual Speech Integration.* Undergraduate honors thesis, The Ohio State University.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189-217.

Li, K. Z. H., & Lindenberger, U. (2002). Relations between aging sensory/ sensorimotor and cognitive functions. *Neuroscience and Biobehavioral Reviews, 26,* 777?783.

MacLeod, A., & Summerfield, Q. (1987). "Quantifying the contribution of vision to speech perception in noise". *British Journal of Audiology, 21,* 131-141.

Martin, J. S. & Jerger, J. F. (2005). Some effects of aging on central auditory processing. *Journal of Rehabilitation Research and Development, 42,* 25?44.

Middelweerd, M., & Plomp, R. (1984). The effect of speech reading on the speech reception threshold of sentences in noise. *Journal of the Acoustical Society of America, 82,* 2145?2147.

Munhall, K.G., Kroos, C., Jozan, C., & Vatikiotis-Bateson, E. (2004). Spatial frequency requirements for audiovisual speech perception. *Perceptions and Psychophysics, 66*(4), 574 ? 583.

Neely, K. K. (1956). Effects of visual factors on intelligibility of speech. *Journal of the Acoustical Society of America, 28,* 1276-1277.

O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech and Hearing Disorders, 19,* 429-439.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17,* 299-320.

Park, D., Smith, A., Lautenschlager, G., & Earles, J. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging, 11,* 621-637.

Pichora-Fuller, M. F., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America, 97,* 593-608.

Plude, D., & Hoyer, W. (1985). Attention and performance: Identifying and localizing age deficits. In N. Charness (Ed.), *Aging and Human Performance*, (pp. 47?99). Chichester, England: Wiley.

Ramakrishna, B. S., Nair, K. K., Chiplunkar, V. N., Ramachandran, V., & Subramanian, R. (1962). *Some aspects of the relative efficiencies of Indian languages.* Ranchi, India: Catholic press.

Ross, L. A., Saint-Amour, D., Leavitt, V., Javitt, D. C. & Foxe J. J. (2007). Do you see what I?m saying? Optimal Visual Enhancement of Speech

Comprehension in noisy environments. *Cerebral Cortex, 17*(5), 1147-53.

Rudmann, D. S., McCarley, J. S., & Kramer, A. F. (2003). Bimodal display augmentation for improved speech comprehension. *Human Factors, 45,* 329-336.

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: from discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology, 56*(3), 139?52.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270,* 303-304.

Shannon, R. V., Zeng, F. G., & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *The Journal of the Acoustical Society of America, 104*(4), 2467-2475.

Shoop, C., & Binnie, C. A. (1979). The effects of age upon the visual perception of speech. *Scandinavian Audiology, 8,* 3?8.

Sommers, M. S. (1997). Speech perception in older adults: The importance of speech-specific cognitive abilities. *Geriatric Bioscience, 45,* 633-637.

Spehar, B., Tye-Murray, N., & Sommers, M. S. (2008). Intra- versus intermodal integration in young and older adults. *Journal of the Acoustical Society of America, 123*(5), 2858?2866.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26,* 212-215.

Summerfield, A. Q. (1979). Use of visual information for phonetic perception. *Phonetica, 36,* 314?331.

Tye-Murray, N., Sommers, M. S., & Spehar, B. (2007). Audiovisual integration and lip-reading abilities of older adults with normal and impaired hearing. *Ear and Hearing, 28*(5), 656-668.

Walden, B. E., Busacco, D. A., & Montgomery, A. A. (1993). Benefit from visual cues in auditory-visual speech recognition by middle-aged and elderly persons. *Journal of Speech and Hearing Research, 36,* 431?436.

Yathiraj, A., and Vijayalakshmi, C. S. (2005). *Phonemically balanced word list in Kannada.* A test developed at Department of Audiology, AIISH, Mysore.

Auditory Plasticity in Musicians: A Comparative Study

[1]Deepika Verma & [2]Rajalakshmi K.

Abstract

Musical training is a rigorous routine involving the segregation of vocal and instrumental sounds presented concurrently. To investigate the effect of musical experience on the neural representation of speech-in-noise, we compared behavioral scores of speech perception in noise and sub cortical neurophysiological responses to speech in a group of highly trained musicians and non-musician controls. A total of 50 participants comprised musicians and non musicians were taken, aged between 7-18 years. Speech evoked ABR and LLR were done along with SPIN test. Musicians were found to have a more robust sub cortical representation of the speech stimulus. Musicians had earlier response onset timing, than non-musicians Fo encoding is better in musicians than in non musicians. Speech perception ability in musicians becomes better with increased years of musical exposure and experience. The increase in Fo amplitude is positively correlated with the speech perception in noise.Musicians had better behavioral performance on the Speech in Noise Test (SPIN) outperforming the non-musician controls. These findings suggest that musical experience limits the negative effects of competing background noise.

Keywords: *Plasticity, Musicians, Speech perception in noise, Speech evoked ABR*

Introduction

From the cochlea to the auditory cortex, sound is encoded at multiple locations along the ascending auditory pathway, eventually leading to conscious perception. Speech is a stream of acoustic elements produced at an astounding average rate of three to six syllables per second (Laver, 1994).The ability to decode these elements in a meaningful manner is a complex task that involves multiple stages of neural processing.

Neural plasticity is a term used to describe alterations in the physiological and anatomical properties of neurons in the brain in association with auditory stimulation and deprivation. Depending on the experience, mechanism of plasticity can involve synaptic changes that occur rapidly or slowly over a period of time (Tremblay & Kraus, 2002). Everyday learning and training involves of continuous improvement of our abilities the sensory, cognitive & behavioural levels (Menning, Roberts & Pantev, 2000). Peripheral and central structures along the auditory pathway contribute to speech processing and learning. However, because speech requires the use of functionally and acoustically complex sounds which necessitates high sensory and cognitive demands, long-term exposure and experience using these sounds is often attributed to the neocortex with little emphasis placed on subcortical structures (Song, Skoe, Wong & Kraus, 2008).

Music is a complex auditory task and musicians spend years fine-tuning their skills. It is no wonder that previous research has documented neuroplasticity to musical sounds as a function of experience (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005; Koelsch, Schroger & Tervaniemi, 1999; Musacchia, Sams, Skoe & Kraus, 2007; Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Tervaniemi, Rytkonen, Schroger, Ilmoniemi & Naatanen, 2001). The domains of music and language share many features, the most direct being that both exploit changes in pitch patterns to convey information. Music uses pitch contours and intervals to communicate melodies and tone centers. Pitch patterns in speech convey prosodic information; listeners use prosodic cues to identify indexical information, i.e., information about the speaker's intention as well as emotion and other social factors.

Musicians have a variety of perceptual and cortical specializations compared to non-musicians. Recent studies have shown that potentials evoked from primarily brainstem structures are enhanced in musicians, compared to non-musicians. Specifically, musicians have more robust representations of pitch periodicity and faster neural timing to sound onset when listening to sounds or both listening to and viewing a speaker. However, it is not known whether musician-related enhancements at the subcortical level are correlated with specializations in the cortex (Musacchia, Strait & Kraus, 2008). The effects of musical experience on the nervous system include relationships between brainstem and cortical Evoked Potentials recorded simultaneously in the same subject to seen and heard speech. Moreover, these relationships were related to behavioural measures of auditory perception and were stronger in the audiovisual condition. This implies that musical training promotes plasticity throughout the auditory and multisensory pathways. This includes encoding mechanisms that are relevant for musical sounds as well as for the processing of linguistic cues and multisensory information (Musacchia et al., 2008).

Hearing speech in noise is a difficult task for everyone, but young children and older adults are particularly vulnerable to the deleterious effects of background noise. Children with learning disorders can ex-

[1]Email: verma.deepika28@gmail.com,
[2]Professor of Audiology, Email: veenasrijaya@gmail.com

hibit noise exclusion as a primary symptom (Sperling, Lu, Manis & Seidenberg, 2005). Musicians, in contrast, demonstrate enhanced noise-exclusion abilities (Parbery-Clark, Skoe & Kraus, 2009a; Parbery-Clark, Skoe, Lam & Kraus, 2009b). Musical experience enhances the ability to hear speech in challenging listening environments. Speech in Noise performance is a complex task requiring perceptual cue detection, stream segmentation, and working memory. Musicians performed better than nonmusicians in conditions where the target and the background noise were presented from the same source, meaning parsing was more reliant on the acoustic cues present in the stream (Parbery & Kraus, 2009).

There is evidence of musical expertise contributing to an enhanced subcortical representation of speech sounds in noise. Musicians had more robust temporal and spectral encoding of the eliciting speech stimulus, thus offsetting the deleterious effects of background noise. Faster neural timing and enhanced harmonic encoding in musicians suggests that musical experience confers an advantage resulting in more precise neural synchrony in the auditory system. These findings provide a biological explanation for musicians' perceptual enhancement for speech-in-noise (Anderson & Kraus, 2010).

Thus, this study was taken up to study the brainstem correlation of speech in noise perception in musicians and to document the auditory plasticity induced by music in musicians on the basis of experience in Carnatic music and also to compare the brainstem and cortical plasticity in musicians and non musicians.

Method

The present study aimed to find out the effect of musical training on auditory plasticity and speech perception in noise in musicians with various years of Carnatic vocal musical training or practice, using Speech evoked Auditory Brainstem Response, Speech evoked Late Latency Response, and Speech Perception in Noise (SPIN) tests.

Participants

A total of 50 participants aged between 7 to 18 years. Twenty children enrolled for Carnatic music learning and 25 untrained children were included in this study. The musicians were classified in to 3 groups.Twenty five trained musicians were classified into 4 groups.

Group 1: Music learning age ranging from 7 to 10 yrs with minimum experience of 2 to 3 years (5 participants).

Group 2: Music learning age ranging from 10 to 13 yrs with minimum experience of 4 to 5 years (10 participants).

Group 3: Music learning age ranging from 13 to 18 yrs with experience of greater than 6 years (10 participants).

Test Stimulus

The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980) which is available in the Biologic Navigator Pro EP system in the BIOMARK protocol. This stimulus simultaneously contains broad spectral and fast temporal information characteristic of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. The fundamental frequency (F0) of the /da/ stimulus linearly rises from 103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 ms. The first formant (F1) rises from 220 to 720 Hz, while the second formant (F2) decreases from 1700 to 1240 Hz over the duration of the stimulus. The third formant (F3) falls slightly from 2580 to 2500 Hz, while the fourth (F4) and fifth formants (F5) remain constant at 3600 and 4500 Hz, respectively.

The stimulus used was the default BIOMARK synthetic /da/ syllable of 40 ms duration produced using KLATT synthesizer (Klatt, 1980). The stimuli were presented at a rate of 10.9/s through ER-3 insert earphones in alternating polarities at an intensity of 80 dB SPL with a filter setting of 100 to 3000 Hz. The responses were collected using three AgCl scalp electrodes. Responses were differentially recorded over a time window of 64 ms (including pre-stimulus period of 11ms) with a vertical montage (Test Ear Mastoid-Active, Forehead-Ground, and Non-Test Ear Mastoid- Reference). The waveforms of the participants, acquired after 1500 artifact free sweeps were weighted added for each ear and then analyzed using the BIOMARK module .Waveforms were collected for rarefaction & condensation polarities and weighted addition was done to obtain calculated waveforms. The speech evoked ABR and FFR waveform, were converted into ASCII format using the software called 'AEP TO ASCII'.

Test Procedure

Routine audiological evaluation was carried out in an acoustically treated room. Air conduction and bone conduction thresholds were established using modified Hughsen and Westlake (Carhart & Jerger, 1959) procedure. Speech audiometry was also carried out on all the participants. A calibrated two-channel Madsen (Orbiter-922) audiometer with TDH-39 headphones was used to establish air conduction pure tone thresholds and speech audiometry. B-71 bone vibrator was used to establish bone conduction thresholds. Hearing was considered normal if puretone thresholds were within 15 dB HL bilaterally at all octave frequencies from 250 Hz to 8 kHz. Tympanometry and Acoustic Reflex Thresholds were established using a calibrated Grason Stadler-Tympstar middle ear analyzer. Presence of 'A' or 'As' type of tympanogram with reflexes present in both the ears below 100 dB HL at 500 Hz, 1 kHz and 2 kHz was considered as normal.

Speech Perception in Noise (SPIN): SPIN test was done using the phonemically balanced (PB) Kannada word list (Yathiraj & Vijayalakshmi, 2005), recorded in the voice of a typical female Kannada speaker. The stimuli were played in a laptop and were routed through the audiometer. The presentation level was 40 dB SL (with reference to SRT) or at most comfortable level. The monosyllables and the speech noise were presented monaurally at two different SNRs (0dB & -5 dB). Twenty five monosyllables were presented for each trial. The participants' task was to perceive the monosyllables presented in the presence of noise and repeat them back. Each word was given a score of 4 %. Number of correctly identified word at different SNRs was noted down to find the SPIN score.

Speech Evoked LLR: Biologic Navigator Pro EP System version 7.0 was used for recording speech evoked late latency responses. LLR was assessed for P1, N1, and N2 in terms of latency.

Speech Evoked Auditory Brainstem Response: Biologic Navigator Pro EP System version 7.0 was used for recording speech evoked auditory brainstem response.

Data Analysis

Speech evoked ABR is composed of the transient and the sustained responses (also known as frequency following responses). Transient response consists of peak V and peak A whereas the sustained responses consist of peaks D, E, F, and O. In the present study latency of both the transient as well as sustained responses were analyzed.

The transient response was analyzed in terms of latency and amplitude of V and A peak. The FFR response was analyzed in terms of latency and amplitude of D, E, F, O peaks for the earlier mentioned three repetition rates (the distance between the peak D, E, F, and O is approximately 10 ms which gives the information regarding the encoding of fundamental frequency). The sustained portion was analyzed using Fast Fourier Transformation (FFT) for the latency range of 11.4 ms to 40.6 ms for speech evoked ABR to extract the information regarding the coding of fundamental frequency, first formant frequency and second formant frequency at different repetition rates using the MATLAB software.

Procedure for FFT analysis: To know the coding of fundamental frequency, first formant frequency and higher harmonics, FFT analysis of the sustained response of the speech evoked ABR was done. This was executed using the MATLAB version 7.0 software (Brainstem toolbox) developed by Skoe and Kraus (2004), at North western university. For measuring the fundamental frequency and higher harmonics, Fourier analysis was performed on the 11.4 to 40.6 ms epoch of the FFR in order to assess the amount of activity occurring over three frequency ranges. Activity occurring in the frequency range of the response corresponding to the fundamental frequency of the speech stimulus (103 to 121 Hz), first formant frequencies of the stimulus (454 to 719 Hz) and for the higher harmonics (721 to 1155 Hz) were measured for all the participants.

Results and Discussion

The data was appropriately tabulated and statistically analyzed using SPSS (version 18) software. Descriptive statistics (mean and standard deviation) were obtained for all the parameters for both ears separately. Separate 2-WAY MANOVA was done to see the significant difference between musician and non-musicians for all the parameter for speech ABR (Latency of wave V, A, C, D, E, F, O), LLR (Latency of P1, N1, P2) and for amplitudes of Fo and F1. Two-way ANOVA was done to see the significant difference between musicians & non-musicians for SPIN scores. Pearson correlation was calculated to see the correlation between SPIN scores and the amplitudes of Fo & F1 in musicians & non-musicians.

Speech Evoked ABR

Latency of Waves V, A, C, D, E, F and O was measured for non-musicians & musicians, for both ears separately, for all three groups. Descriptive statistics (mean & SD) was done for latencies of speech evoked ABR waves for musicians & non-musicians across three groups.

Two-way MANOVA was done to see the differences between musicians and non-musicians for latencies of speech evoked ABR. There was a significant difference between non-musicians & musicians ($p<0.05$) for latency of wave V & O for all three groups. There was no significant difference present for latencies of other waves between musicians and non-musicians.

In the present study, the latency of wave V and O responses were significantly different between musicians and non-musicians. There was no significant difference present in transition latencies between musicians and non-musicians. These results are in agreement with the study by Parbery-Clark et al. (2009a), where it was concluded that musicians had earlier response onset timing, than non-musicians. Musacchia et al. (2008) reported that latency and amplitude of wave V differed between musicians and non-musicians. The results of the study by Parbery-Clark et al. (2009a) also suggested that there was no significant difference in transition latencies between musician and non musician in quiet conditions.

Table 1: Mean and SD for FFT- Fast Fourier Transform; Fo, F1 & F2- Fundamental frequency, first and second Formants for non-musicians

Wave	Age Group (in years)	Mean (dB)	Standard Deviation
	7 to10	2.076	0.645
Fo	10 to13	3.379	1.308
	13 to18	4.062	1.204
	7 to10	1.412	0.368
F1	10 to13	1.286	0.343
	13 to18	1.252	0.434
	7 to10	.5392	0.095
F2	10 to13	0.447	0.115
	13 to18	0.460	0.162

Table 2: Mean and SD for FFT- Fast Fourier Transform; F0, F1 & F2- Fundamental frequency, first and second formants for musicians

Wave	Age Group (in years)	Mean (dB)	Standard Deviation
	7 to10	2.303	0.669
Fo	10 to 13	4.975	0.971
	13 to 18	5.554	1.196
	7 to 10	1.326	0.556
F1	10 to 13	1.518	0.475
	13 to 18	1.389	0.381
	7 to 10	0.482	0.153
F2	10 to 13	0.523	0.136
	13 to 18	0.527	0.149

Musicians and non-musicians had equivalent stimulus response correlation in quiet. Their results are in agreement with the results of the present study.

Speech Evoked Late Latency Response

Latency of Waves P1, N1& P2 was measured for non-musicians & musicians, for both ears separately. Descriptive statistics (mean & SD) was done for latencies of speech evoked LLR waves across all age groups.

Two-way MANOVA was done to see the differences between musicians and non-musicians for latencies of speech evoked LLR waves across years of musical experience There is no significant difference between musicians & non-musicians in terms of latencies of wave P1, N1 and P2 of speech evoked LLR ($p < 0.05$).The results of the present study are in consonance with the results by Strait et al. (2011). There was no response variability among musicians and non-musicians at any electrode site.

Shahin et al. (2003) reported enhanced P2 and N1c responses in musicians compared to non-musicians. Krista et al. (2009) reported that long term music training offers structural plasticity in developing correlation with behavioural changes. T1 weighted MRI was used in their study for assessment.

The difference in the results of the present study with the earlier studies reported in the literature can be accounted on the following reason: First, the assessing tool used in previous studies is magneto encephalography (MEG), electroencephalography (EEG) and MRI. These radiological tests are different from far field electrophysiological responses. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on Carnatic vocal musicians. Moreover, the participants taken in Shahin et al. (2003) study were having more years of musical experience (greater than 11 years) than the participants of the present study.

FFT- Fast Fourier Transform: F0, F1 and F2

Fo, F1 and F2- Fundamental frequency, first and second formants were measured for non-musicians & musicians, for both ears separately, for all three groups. Table 1 and Table 2 show descriptive statistics (mean & SD) for FFT- Fast Fourier Transform; Fo, F1 and F2- Fundamental frequency, first and second formants for non-musicians and musicians.

Two-way MANOVA was done to see the differences between musicians and non-musicians for FFT- Fast Fourier Transform; F0, F1 & F2- Fundamental frequency, first and second Formants (Table 3).

There was a significant difference present ($p < 0.05$) for Fo Formant between musicians and non-musicians for group 3. The amplitude of energy concentration in Fo formant is significantly larger in Group 3 compared to non-musician. The results of the present study are in agreement with results of Mussachia et al. (2007). They reported that musicians have larger response amplitudes for encoding of speech and music stimuli compared to non-musicians. Mussachia et al. (2008) reported experienced musicians had larger Fo peak amplitudes. In the present study, musicians in Group 3, with higher years of musical experience had better mean SPIN scores than non-musicians. This result draws support study by Anderson et al (2010), which suggests that good SPIN perceivers had greater spectral magnitudes for Fo and H2.

Table 3: Statistical Results for Two way Manova for FFT

Source	Variable	df	F	p
grpNM*M	Fo	2	2.910	0.000*
	F1	2	0.934	0.397
	F2	2	1.689	0.190

Table 4: Mean and SD for SPIN for all groups

Group	NM/M	Mean (%)	Standard Deviation
1	Non musician	72.80	3.676
	Musician	73.20	4.237
2	Non musician	77.00	4.657
	Musician	80.00	4.768
3	Non musician	80.80	3.488
	Musician	86.00	2.865

The difference in the results for groups 1 and 2 of the present study can be explained on the following reasons; first, the musical year of experience in group 1 and 2 were less than reported in the previous studies. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on vocal musicians. Moreover, the participants taken in Mussachia et al. (2008) study were having more experience than the participants in the present study. Most of the study reports experience of 10 for their participants.

Speech Perception in Noise

The speech perception in noise was assessed for all the 50 participants for both the ears. The test was carried out at 0 dB SNR.

2-WAY ANOVA was employed to see the significant difference between musicians & non-musicians. There was a significant difference across the three groups *(p*< 0.05). However, there was no significant difference between musicians & non-musicians across the different groups for SPIN.

On comparing mean score (Table 4) the mean scores are quite better for musicians in group 3, but not significantly better. But this is in contrast to the previous research done on speech perception abilities in musicians. According to a study done by Parbery-Clark et al. (2009a), musical experience enhances the ability to hear speech in challenging listening environments. In another study Parbery-Clark et al. (2009b) found that musical experience resulted in more robust subcortical representation of speech in the presence of background noise. The difference in the results of the present study with the earlier studies reported in the literature can be accounted on the following reasons: First, the noise used in the previous studies were speech shaped noise or multi-talker babble. But in the present study speech noise was used to study the speech perception in noise. It is evident that the speech shaped noise or multi-talker babble will give better results for speech perception in noise when compared to speech noise. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on vocal musicians. Moreover, the participants taken in Parbery-Clark et al. (2009a) study were having more experience than the participants in the present study. Third, the speech material used in previous studies was sentences (Quick SIN, HINT). The sentences are more redundant than words. Fourth, the present study was conducted at 0 dB SNR, the studies reported in literature suggested that SPIN is better in adverse listening conditions. Parbery-Clark et al. (2009b) reported musicians were able to repeat sentences presented at a lower, more challenging SNR than non-musicians.

Correlation between SPIN & FFT

The results revealed that there is a positive and highly significant correlation between SPIN and Fo. With the increase in Fo amplitude, there is an increase in the scores for the perception of speech in noise. This finding can be supported by study of Anderson et al. (2010), good SIN perceivers have greater spectral magnitudes for Fo and H2.

Table 5: Correlation between FFT and Speech Perception in Noise

Formants	SPIN
Fo	0.000**
F1	0.826
F2	0.914

There was no significant correlation between other formants & SPIN scores. As the Formants increases, the amplitude of the harmonics decreases. Thus the results of the present study reveal that latencies of wave V and O were significantly different between musicians and non musicians.

The latencies of speech evoked LLR was did not reveal statistically significant difference for musician and non musician. A significant difference was noticed between musician and non musician for Fo formant. There is a significant difference across the three groups of musicians on SPIN score. A highly significant positive correlation is present between Fo and SPIN scores.

Conclusions

Fo encoding is better in musicians than in non musicians. Speech perception ability in musicians becomes better with increased years of musical exposure and experience. The increase in Fo amplitude is positively correlated with the speech perception in noise. There was no significant difference in latencies of P1, N1 and N2 between musicians and non musicians for speech evoked LLR.

It can be implemented in hearing aid technology for musicians with hearing loss to improve their speech perception. Music training can be used as a potential remediation strategy for children requiring language training and auditory processing disorders with noise exclusion deficits.

Future research on clinical population who may exhibit neural encoding deficits such as autism. Brainstem maturation as an indication in infants & preschool children at risk.

The present study can b replicated across vocal musicians and instrumental musicians. It can be compared between Hindustani and Carnatic musicians. Musicians and dancers can be compared to find whether there are differences in Fo encoding and ability to perceive speech in the presence of noise.

References

Anderson, S., & Kraus. N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review (2010). *Journal of American Academy of Audiology, 21*, 575-585.

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

Fujioka, T., Trainor, L., Ross, B., & Kakigi, R. (2004). Musical Training Enhances Automatic encoding of Melodic Contour and Interval Structure. *Journal of cognitive Neuroscience, 16*, 1010-1021.

Klatt, D. (1980). Software for a Cascade/Parallel Formant Synthesizer. *Journal of the Acoustical Society of America, 67,* 13-33.

Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport, 10*, 1309-1313.

Laver, J. (1994). *Principles of phonetics.* Cambridge, UK: Cambridge University Press.

Menning, H., Roberts, L., & Pantev, C., (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Auditory and Vestibular Systems, 11*, 817- 822.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of United States of America, 104*, 15894-15898. Musacchia, G., Strait, D., & Kraus, N. (2008) Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research, 241,* 34-42.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., Almut., Ross., & Bernhard. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport, 12*, 169-174.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience, 29,* 14100-14107.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009b). Musician enhancement for speech in noise. *Ear and Hearing, 30*, 653-661.

Shahin , A., Bosnyak , D., Trainor, L., Roberts, & Larrey, R., (2003). Enhancement of neurolplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience, 12*, 5545-5552.

Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing, 31(*3), 302-324.

Song, J, H., Skoe, E, Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognition Neuroscience, 20,* 1892-1902.

Sperling, A. J., Lu, Z., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8*, 862-863.

Tervaniemi, M., Rytkonen, M., Schroger, E., Ilmoniemi, R. J, & Naatanen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory, 8,* 295-300.

Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortica; neural activity. *Journal of Speech-Language-Hearing Research, 43*(3), 564-72.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonemically Balanced word list in Kannada*. Developed in Department of Audiology, AIISH, Mysore.

Perception of Spectrally Enhanced Speech through Companding in Individuals with Cochlear Hearing Loss

[1]Deepthi M. & [2]Vijaya Kumar Narne

Abstract

The present study aimed to assess the benefit of spectral contrast enhancement using companding strategy in individuals with cochlear hearing loss. Twenty adult subjects having normal hearing and cochlear hearing loss participated in the study. Consonant recognition scores (quiet, 15 dB SNR, 10 dB SNR and 0 dB SNR) and sentence recognition threshold in noise was obtained in unprocessed (original) and processed (spectrally enhanced using companding strategy) condition. Numbers of correctly recognised consonants were calculated for the consonant recognition task. The SNR level at which 50 % of the correct responses were obtained using sentences (SNR-50), across different SNRs (+20 dB to -10 dB SNR) was found. Results showed that processed condition of consonant recognition scores lead to higher performance than unprocessed in both the groups. Significant improvement was found at 0 dB SNR (12 %) for normal hearing individuals and at 15 dB (4.5 %), 10 dB (5.25 %) and 0 dB SNR (18.75 %) for those with cochlear hearing loss. In sentence recognition threshold in noise (SNR-50) task, both the subjects performed at lower SNR levels in processed than unprocessed condition. Improvement found was about -3.75 dB SNR for normal hearing individuals and -5 dB SNR for those with cochlear hearing loss. Thus, it can be concluded that spectrally enhanced speech through companding strategy improves the speech perception in noise in individuals with cochlear hearing loss to a much greater extent than do for normal hearing individuals.

Keywords: *Spectral enhancement, Companding, Speech Perception, Cochlear hearing loss*

Introduction

The most common type of hearing loss is sensorineural hearing loss, which is usually associated with a dysfunction of the cochlea. People with cochlear hearing loss can understand speech reasonably well in one-to-one conversation in a quiet room, but they have great difficulty when there is background noise or reverberation, or when more than one person is talking (Plomp, 1978).

Reduced frequency selectivity is a well-documented abnormality that is associated with cochlear hearing loss, which can affect speech perception in noise (Tyler, Wood &Fernandes, 1982; Preminger & Wiley, 1985; Thibodeau& van Tasell, 1987). One mechanism by which impaired frequency selectivity could affect speech understanding in noise involves the perception of spectral shape. The recognition of speech sounds requires a determination of their spectral shapes, especially the locations of spectral prominence. Broader auditory filters associated with cochlear hearing loss, produce a more highly smoothened representation of the spectrum. If spectral features are not sufficiently prominent, they may be smoothed to such an extent that they become imperceptible. Leek, Dorman and Summerfield (1987) reported that the greater spectral contrast was required for vowel identification by hearing impaired than for normal hearing listeners. Adding a noise to speech fills the valleys between the spectral peaks and thereby reduces spectral prominence, resulting in poorer perception of speech in the presence of noise.

Thus, improving the intelligibility of speech in noise for individuals with cochlear hearing loss is one of the most difficult tasks faced by hearing aid manufacturers. There are currently a variety of tools available for this task, which includes the application of digital signal processing to hearing aids. With appropriate prescription and fitting, a hearing aid can significantly improve speech recognition for an individual with hearing impairment in quiet and non-reverberant listening environment. However, this benefit is greatly reduced in the presence of noise (Killion&Niquette, 2000). Hence, one of the challenges in providing amplification for the cochlear hearing loss individuals is to select the technology that will provide the maximum benefit in the presence of noise.

If reduced frequency selectivity impairs speech perception in noise for individuals with cochlear hearing loss, then enhancement of spectral contrasts might improve their performance. A number of spectral enhancement techniques have been tested using normal hearing and hearing-impaired) listeners in order to improve their speech understanding in noise (Bunnell, 1990; Clarkson &Bahgat, 1991) and small to modest benefits have been obtained with the signal enhancement (Baer, Moore & Gatehouse, 1993). Recently, Turicchia and Sarpeshkar (2005) applied a frequency-specific companding strategy for spectral contrast enhancement and showed that it has the potential to improve speech performance in noise in cochlear implant users. Similarly, Bhattacharya and Zeng (2006) studied speech recognition in speech-shaped noise by cochlear implant users using companding strategy. They found significant improvement in the recognition of phonemes, consonants and sentences

[1]Email: deepthiaud@gmail.com,
[2]Lecturer in Audiology, Email: vijaynarne@gmail.com

in noise. However, there is a dearth of studies done on investigating the perception of spectrally enhanced speech stimuli using companding strategy in individuals with cochlear hearing loss. Therefore, the current study was conducted.

The aim of the study was to assess the benefit of spectral enhancement using compadingstrategy by comparing thespeech recognition performance in unprocessed (unmodified) and processed (spectrally enhanced) condition among individuals with normal hearingandcochlear hearing loss.

Method

Participants

The data were collected from a total of 40 participants. All of them were native speakers of Kannada.The participants did not have any psychological and neurological problems. They did not have of middle ear pathology as confirmed by immittance evaluation.The participants were divided into two groups; Group I and Group II. Both the groups consisted of 10 male and 10 females.Group I consisted of 20 individualswith normal hearing in the age range 19 to 48 years. Group II included 20 individuals with cochlear hearing loss of age ranging from 18 to 48 years. They had mild to moderately severe degree of hearing loss.

Instrumentation

A calibrated two-channel diagnostic audiometer (GSI-61) with TDH-39 headphones housed in MX-41/AR ear cushions and a bone vibrator, Radio ear B-71 was used to carry out pure tone audiometry. A calibrated GSI-Tympstar (version 2) immittance meter was used to rule out middle ear pathology,ILO 292 DPEcho port system to assess outer hair cell functioning and Intelligent Hearing Systems (IHS smart EP windows USB version 3.91) to rule out retrocochlear pathology. Toshiba Satellite L645 laptop (Realtek sound card) and AHUJA AUD-101XLR dynamic unidirectional was used to record the speech stimulus.MATLAB-7 (Language of Technical computing, USA) was used to spectrally enhance the speech signal using companding strategy.

Test Material

Consonants: Twenty consonants /p, b, t̪, d̪, k, g, ʈ, ɖ, m, n, tʃ dʒ, s, ʃ, j, v, r, l, ɭ, h/ in the context of the vowel /a/ was used. They were spoken by a native Kannada female speaker (language spoken in southern part of India) and were digitally recorded in an acoustically treated room, on a data acquisition system using 44.1 kHz sampling frequency with a 16-bit analog to digital converter. While recording the microphone was placed at a distance of 15cm from the lips of the speaker. The

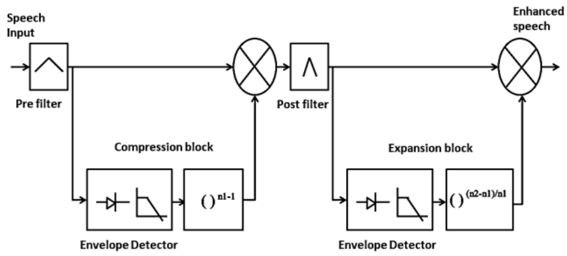


Figure 1: Design of a single channel companding pathway.

recorded stimuli were root mean square normalized to maintain equal loudness.Goodness test for the recorded material was carried out by presenting the stimuli to 10 individuals with normal hearingto assess quality of recording. All the normal hearing participants obtained 100% score indicating that speech material was highly intelligible. In the experiments involving background noise, each consonant was mixed with a speech spectrum shaped noise at SNRs of 0, 10 and 15 dB.

Sentence: The speech stimuli were sentences in Kannada, developed by Avinash, Raksha and Kumar (2008). There were a total of 10 lists, each list consisting of 7 sentences. Each sentence consisted of 5 target words. All the sentence lists were phonetically balanced and were equally difficult. Each list was mixed with speech spectrum shaped noise at different SNR ranging from +20 dB to -10 dB SNR in 5 dB step-size.

Signal processing strategy: The spectral enhancement using companding architecture was implemented in MATLAB. The companding architecture divides the input signal into 40 frequency channels by a bank of relatively broad band-pass filters. Figure 1 shows the design of a single channel companding pathway. Each channel consists of broad pre-filter, a compression block, a relatively narrow-band post-filter and an expansion block. The time constant of the envelope detector governs the dynamics of the compression or expansion. The extent of compression within each channel depended on the output of ED and compression index (n1). Further, the compressed signal was passed through a relatively narrow band-pass filter before being expanded. The gain of the expansion block depended on the corresponding ED output and the ratio n2/n1. The outputs from all the channels were summed to obtain the processed signal.

Here, 40 channels logarithmically spaced between 100 and 10000 Hz with n1 = 0.3 and n2 = 1 was used. Both consonants and sentences were processed through this companding strategy, to increase the spectral contrast in quiet and different SNR conditions.

Procedure

Speech recognition experiments were done on individuals with normal hearing and cochlear hearing loss. The output from the laptop was routed to the tape input/auxiliary input of the audiometer (GSI-61). Prior to the presentation of the speech stimuli, a 1 kHz calibration tone was played to set the VU meter deflection to '0'. The test stimulus was presented to the individuals at their most comfortable level through the TDH 39 headphones.

Consonant Recognition: In consonant recognition tests, twenty VCV stimuli were presented. They were randomly presented across four different listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR) in unprocessed and processed condition. Subjects were instructed to repeat the consonant that was heard.

Speech Recognition Threshold in Noise: The participants were instructed to listen to the sentence and repeat aloud as many of the words as possible. The experimenter noted the number of words that were correctly repeated by the participant. Stimuli were presented at comfortable level. The starting SNR was +20 dB which was lowered by 5 dB till the level at which two of the four or three of the five words of the sentence are repeated correctly. The SNR at which two of four or three of five words were repeated correctly, was considered as SNR-50.

Results

The data obtained was tabulated and analyzed using Statistical Package for Social Sciences (version 16.0).This was examined for consonant identification and sentence recognition threshold in noise (SNR-50).

Consonant Recognition in Unprocessed Condition

Consonant recognition scores were obtained in unprocessed condition among individuals with normal hearing and cochlear hearing loss. Individuals with normal hearing achieved 95 - 100 % consonant recognition scores in the quiet condition than in the presence of noise. Across different SNRs, maximum scores were obtained at 15 dB SNR, followed by 10 dB and 0 dB SNR. Performance reduced with the decrease in the SNR. Individuals with cochlear hearing loss obtained relatively poorer scores than those with normal hearing as shown in Figure 2. In quiet condition, identification scores obtained were 78 % whereas at 0 dB SNR, scores dropped to 20 %.

Repeated measure ANOVA was performed to assess the difference in unprocessed consonant recognition scores across the four listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR) within individuals with normal hearing and cochlear hearing loss separately. Analysis revealed a significant difference in individuals with normal hearing [$F_{(3,57)} = 300.03$, $P<0.001$] and also cochlear hearing loss [$F_{(3,57)} = 122.17$, $P<0.001$]. For both the groups, there was significant difference between consonant recognition scores in the four listening conditions.

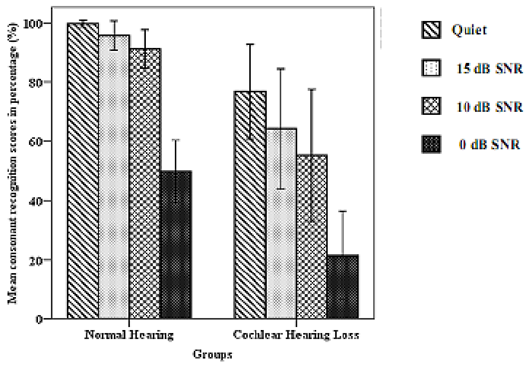


Figure 2: Mean and standard deviation (SD) of unprocessed Consonant recognition scores in normal hearing and cochlear hearing loss group. Error bars indicate SD.

The pairwise comparison was performed using Bonferroni test in both the groups. Results showed that there was significant difference across the different listening conditions for both the groups ($p< 0.05$).

Further, MANOVA was done to compare the unprocessed consonant recognition scores between individuals with normal hearing and cochlear hearing loss across all the four listening conditions. Results showed a significant difference in consonant recognition scores between the groups in quiet [$F_{(1,38)} = 43.56$, $p< 0.05$], 15 dB SNR [$F_{(1,38)} = 45.55$, $p< 0.05$], 10 dB SNR [$F_{(1,38)} = 48.43$, $p< 0.05$] and 0 dB SNR [$F_{(1,38)} = 47.54$, $p< 0.05$].

From the results of present study, it can be noted the normal hearing individuals obtained 100 % consonant recognition scores in quiet condition. However as the SNR decreased, there was a minimal reduction in scores at 15 dB and 10 dB SNR, whereas at 0 dB SNR scores reduced drastically. However, in individuals with cochlear hearing loss, poorer scores were obtained in quiet condition compared to normal hearing individuals. Also as the SNR reduced, there was a drastic decrease in the scores for those with cochlear hearing loss. This reduction in scores was much greater than the normal hearing individuals.

Consonant recognition in unprocessed versus processed condition

Consonant recognition scores were obtained for normal hearing and cochlear hearing loss in both unprocessed and processed condition. Individuals with normal hearing obtained almost similar scores in quiet, 15 dB and 10 dB SNR in both unprocessed and processed condition, whereas scores improved by 12 % at 0 dB SNR in processed condition. Among cochlear hearing loss individuals, improvement in processed condition was about 4.5 % at 15 dB, 5.25 % at 10 dB SNR and 18.75 % at 0 dB SNR. Both the groups obtained higher scores in processed than unprocessed condition as shown in Figure 3 and 4. In addition, individuals with cochlear hearing

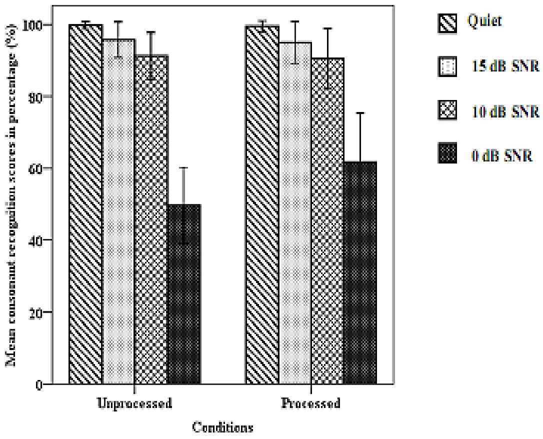


Figure 3: Paired't' test results in unprocessed and processed condition in normal hearing individuals. Error bars indicate SD.

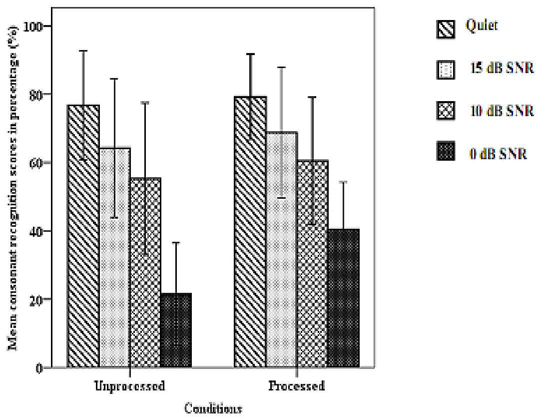


Figure 4: Paired 't' test results in unprocessed and processed condition in cochlear hearing loss individuals. Error bars indicate SD.

loss obtained lesser scores than those with normal hearing.

Repeated measure ANOVA was done to compare the unprocessed and processed consonant recognition scores across listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR) in normal hearing individuals. Further, results showed significant difference across different listening conditions in unprocessed [$F_{(3,57)} = 300.032$, $p < 0.05$] and processed condition [$F_{(3,57)} = 120.159$, $p < 0.05$]. Further, paired't' test was performed to compare unprocessed and processed consonant recognition scores across each of the listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR). Results revealed significant difference in the performance of normal hearing group in unprocessed and processed condition at 0 dB SNR only. Processed condition had an average of 12 % higher scores at 0 dB SNR ($p < 0.05$) than unprocessed condition. However, no significant difference was obtained in quiet ($p = 0.33$), 15 dB ($p = 0.57$) and 10 dB SNR ($p = 0.67$).

To compare the unprocessed and processed consonant recognition scores across listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR) in cochlear hearing loss individuals, repeated measure ANOVA was carried out. Results showed significant difference across different listening conditions in unprocessed [$F_{(3,57)} = 122.178$, $p < 0.05$] and processed [$F_{(3,57)} = 84.548$, $p < 0.05$].

Among cochlear hearing loss individuals, paired't' test results revealed significant difference in unprocessed and processed consonant recognition scores at 15 dB, 10 dB and 0 dB SNR except quiet condition (Figure 4). Maximum improvement was obtained at 0 dB SNR (18.75 %) than 10 dB SNR (5.25 %) followed by 15 dB SNR (4.5 %) in processed over unprocessed condition.

Further, MANOVA was done to compare the consonant recognition scores between normal hearing and cochlear hearing loss individuals across different listening conditions (quiet, 15 dB, 10 dB and 0 dB SNR). Results revealed a significant difference in consonant recognition scores between the groups in quiet [$F(1, 38) = 51.790$, $p < 0.05$], 15 dB SNR [$F(1, 38) = 34.481$, $p < 0.05$], 10 dB SNR [$F(1, 38) = 43.622$, $p < 0.05$] and 0 dB SNR [$F(1, 38) = 24.270$, $p < 0.05$].

Sentence Recognition Threshold in Noise (Snr-50)

Sentence recognition threshold in noise (SNR-50) was obtained in both unprocessed and processed condition among individuals with normal hearing and cochlear hearing loss. Both the groups obtained lower SNR values in processed than unprocessed condition as shown in Figure 5.

To analyze whether mean differences between con-

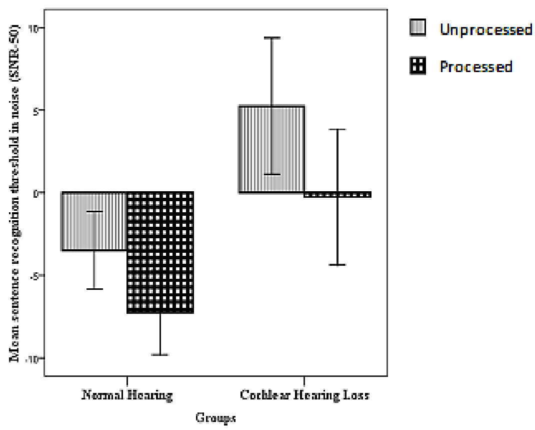


Figure 5: Mean and standard deviation (SD) of unprocessed and processed sentence recognition threshold in noise (SNR-50) in normal hearing and cochlear hearing loss group. Error bars indicate SD.

Table 1: Paired't' test results for sentence recognition threshold in noise (SNR-50) in normal hearing and cochlear hearing loss individuals.

Subjects	Condition	Mean	SD	t value
Normal hearing	Unprocessed	-3.50	2.35	7.55*
	Processed	-7.25	2.55	
Cochlear hearing loss	Unprocessed	+5.25	4.12	15.98*
	Processed	-0.25	4.12	

$*p < 0.05$

ditions for both the groups reached significance, MANOVA was performed. Analysis revealed significant difference between both the groups in unprocessed [F (1, 38) = 67.85, P<0.05] and processed condition [F (1, 38) = 41.609, p< 0.05]. A comparison across the groups indicated that sentence recognition threshold in noise (SNR-50) for the group with normal hearing was lower than the group with cochlear hearing loss in both the conditions.

Sentence recognition threshold in noise (SNR-50) was compared between unprocessed and processed condition in both normal hearing individuals and cochlear hearing loss individuals using paired sample t-test. The results of the paired 't' test is given in Table 1.

From Table 1, it can be described that mean sentence recognition threshold in noise (SNR-50) is significantly lower in processed condition than in non-processed condition for both the groups. In the processed condition, normal hearing individuals are able to perform at an average of -3 dB to -4 dB lower SNR levels compared to unprocessed condition, whereas cochlear hearing loss individuals were able to perform at an average of -5 dB lower SNR levels than unprocessed condition.

Discussion

Consonant Recognition in Unprocessed Condition

In quiet condition, cochlear hearing loss subject's scores were 20 % lower than the normal hearing listeners. The lower consonant recognition scores in cochlear hearing loss individuals may be due to the reduced audibility, reduced frequency selectivity or loudness recruitment. Reduced audibility may not be the major factor, because, the recognition scores were obtained at comfortable level in all the listeners. Many investigators demonstrated no significant correlation between identification scores and frequency selectivity (Dubno& Schaefer, 1992). However, loudness recruitment may be one of the major causes which leads to reduced dynamic range and changes the amplitude variations in the speech signal. These changes involve increase in the amplitude of vowel more significantly than the consonants which increase the upward spread of masking. This leads to masking of consonantal portion, and hence consonant recognition is affected in individuals with cochlear hearing loss.

In noisy condition, consonant recognition scores in normal hearing individuals reduced more significantly at 0 dB SNR by about 40 %, whereas in cochlear hearing loss scores dropped to almost 16 % at 0 dB SNR. The precise reason for low scores is not known. Some of the possible reasons could be reduced frequency selectivity and loudness recruitment. Investigators have demonstrated that individuals with cochlear hearing loss have auditory filters that are broader than normal (Glasberg& Moore, 1986; Tyler et al., 1982). This means that, the ability to resolve the spectral components of speech sounds and to separate the components of speech from background noise is reduced. One mechanism by which impaired frequency selectivity could affect speech understanding in noise involves the perception of spectral shape. Broader auditory filters produce a more highly smoothed representation of the spectrum (the excitation pattern) than the normal auditory filter. Further, smoothened spectrum results in reduced formant frequency representation which causes imperceptions of the formants. Adding a noise background to speech, fills the valleys between the spectral peaks and thus reduces spectral prominence, exacerbating the problem of perceiving them for people with broadened auditory filter. Another reason is that, many recent investigators have demonstrated that cochlear hearing loss listeners depend more on envelop of speech signal than the fine structure and adding a noise would significantly alter the envelop of the signal that is., reducing the modulation depth and distorting the modulation. Because of the above mentioned reasons, individuals with cochlear hearing loss have more significant problem than those with normal hearing.

To summarize, individuals with cochlear hearing loss perform poorly in noise which may be due to reduced frequency selectivity, loudness recruitment and impaired ability in extracting envelop of speech signal in noisy condition.

Consonant Recognition in Unprocessed Versus Processed Condition

In the present study, spectral enhancement using companding strategy improved the consonant recognition scores in noise for individuals with normal hearing and cochlear hearing loss. To our knowledge, there are no studies that have utilized companding strategy in cochlear hearing loss individuals. Many studies which have used various other strategies to enhance spectral contrast have shown improvement in noise with enhancement (Baer et al., 1993; Yang et al., 2003; Frank et al, 1999). However, the above mentioned studies cannot be compared due to the larger differences in the signal enhancing strategies and rationale behind these strategies. The improvement with companding strategy can be because of two reasons: (i) reduced frequency selectivity affects the perception of the consonant in the presence of noise by reducing its spectral contrast. Increasing the spectral contrast of the consonant using spectral enhancement, thereby will compensate for reduced frequency selectivity and reduced spectral contrast (Baer et al., 1993; Watkins & Makin, 1996); (ii) envelop of a less intense consonants can be masked by high intense vowels resulting in the degradation of the envelop, leading to imperceptions of that particular consonant (Brokx&Nooteboom 1982; Turner, Souza & Forget, 1995). However, enhancing envelop of consonant prevent it from the upward spread of masking caused by vowels, due to increased amplitude of consonant portion. Because of the above reasons, cochlear hearing loss individuals perform better with processed recognition condition.

Sentence recognition threshold in noise (SNR-50)

In the present study, individuals with cochlear hearing loss required +7 dB higher SNR for sentence recognition threshold in noise (SNR-50) than the normal hearing individuals. These results are in agreement with those of previous studies (Plomp, 1994; Needleman &Crandell, 1995; Bacon et al., 1998). The reason for obtaining higher SNRs in individuals with cochlear hearing loss may be due to broader auditory filters, which degrades the spectrum of the speech signal (Glasberg& Moore, 1986; Tyler, et al., 1982). In addition to this, adding background noise further reduces the spectral prominence in the speech signal. Also, because of the loudness recruitment, speech sound with maximum energy can mask out the other speech sounds which are less intense. As a result, envelop of speech signal would be distorted which can result in reduced modulation depth. This can further impair the speech perception when an additional background noise is added to it. Because of the above mentioned reasons, individuals with cochlear hearing loss have more significant problem than those with normal hearing.

Using the spectral enhancement through companding strategy, both the groups obtained lower SNRs in processed than unprocessed condition. To our knowledge, many other studies (Baer et al., 1993; Yang et al., 2003; Franck et al, 1999) have also obtained similar findings using different strategies. But, these studies cannot be directly compared with the present study, due to the larger differences in the signal enhancing strategies employed.

In the unprocessed condition, the speech signal will be degraded in the presence of noise, making the listeners more difficult to identify the words in the sentences. This is because of reduced spectral contrast (Moore &Glasberg, 1983; Leek et al., 1987) and distorted envelop of the speech signal (Brokx&Nooteboom 1982; Turner et al., 1995). Bhattacharya and Zeng (2007) have shown that spectral contrast in the processed signal significantly enhanced compared to unprocessed condition. The improvement observed for CI individuals in their study is attributed to increased spectral contrast. Similarly, Oxenham et al. (2007) have demonstrated similar results. In the present study improvement observed in the processed condition, can be attributed to enhanced spectral contrast. In addition, the companding strategy also enhances the envelope of the signal which would have enhanced the less intense speech sounds, preventing it from the upward spread of masking by high intense vowels. Hence, subjects obtained lower SNRs in the processed than unprocessed condition.

To summarize, spectral enhancement improved consonant and sentence perception for both the individuals with normal hearing and cochlear hearing loss. The amount of improvement observed was higher for cochlear hearing loss than normal hearing listeners.

Conclusions

The major findings of the study indicated that spectrally enhanced speech through companding strategy improved the speech perception in noise among individuals with cochlear hearing loss to a much greater extent than do for normal hearing. The important clinical implication of the current study is that, the spectral enhancement using companding strategy has the potential to improve speech performance in the presence of noise among those with cochlear hearing loss individuals. Further, this strategy can be implemented in amplification devices for the benefit of speech recognition in adverse listening conditions.

References

Bacon, S. P., Opie, J. M., & Montoya, D.Y. (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech, Language, and Hearing Research, 41*, 549-563.

Baer, T., Moore, B. C. J., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times. *Journal of Rehabilitation Research and Development, 30,* 49-72.

Bhattacharya, A., & Zeng, F. (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noise. *Journal of the Acoustical Society of America, 122,* 1079-1089.

Brokx, J. P. L., & Nooteboom, S. G. (1982). Intonation and the perception of simultaneous voices.

Journal of Phonetics, 10, 23-26.

Bunnell, H. T. (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *Journal of the Acoustical Society of America,88,* 2546-2556.

Clarkson, P. M., & Bahgat, S. F. (1991). Envelope expansion methods for speech enhancement. *Journal of the Acoustical Society of America, 89,* 1378-1382.

Dubno, J. R., & Schaefer, A. B. (1992) Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners. *Journal of the Acoustical Society of America, 91,* 2110-2121.

Franck, B. A., van Kreveld-Bos, C. S., Dreschler, W. A., & Verschuure, H. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *Journal of the Acoustical Society of America, 106,* 1452-1464.

Glasberg, B. R., & Moore, B. C.J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *Journal of the Acoustical Society of America, 79,* 1020-1033.

Killion, M., & Niquette, P. (2000). What can the pure tone audiogram tell us about a patient's SNR loss?.*The Hearing Journal, 53,* 46-53.

Leek, M. R., Dorman, M. F., & Summerfield, Q. (1987). Minimum spectral contrast for vowel identification by normalhearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 81,* 148-54.

Moore, B. C. J., & Glasberg, B. R. (1983). Masking patterns of synthetic vowels in simultaneous and forward masking. *Journal of the Acoustical Society of America, 73*, 906-917.

Needleman, A.R. & Crandell, C.C. (1995). Speech recognition in noise by hearing-impaired and noise-masked normal-hearing listeners. *Journal of the American Academy of Audiology, 6,* 414-424. Oxenham, A. J., Simonson, A. M., Turicchia, L., & Sarpeshkar, R. (2007). Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing. *Journal of the Acoustical Society of America,121,* 1709-1716.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *Journal of the Acoustical Society of America, 63,* 533-549.

Plomp, R. (1994). Noise, amplification, and compression: Considerations of 3 main issues in hearing-aid design. *Ear and Hearing,15,* 2-12.

Preminger, J. & Wiley, T.L. (1985). Frequency selectivity and consonant intelligibility in sensorineural hearing loss. *Journal of Speech and Hearing Research, 28,* 197-206.

Thibodeau, L.M., & van Tasell, D.J. (1987). Tone detection and synthetic speech discrimination in band-reject noise by hearing-impaired listeners. *Journal of the Acoustical Society of America, 82,* 864-873.

Turicchia, L., &Sarpeshkar, R. (2005). A bio-inspired companding strategy for spectral enhancement. *IEEE Transactions on Acoustics Speech and Signal Processing. 13*, 243-253.

Turner, C.W., Souza, P.E., & Forget, L.N. (1995). Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *Journal of the Acoustical Society of America, 97*, 2568-2576.

Tyler, R.S., Wood, E.J., & Fernandes, M.A. (1982). Frequency resolution and hearing loss. *British Journal of Audiology, 16*, 45-63.

Watkins, A.J., & Makin, S.J. (1996). Effects of spectral contrast on perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America,99,* 3749-3757.

Yang, J., Luo, F., & Nehorai, A. (2003). Spectral contrast enhancement: algorithms and comparisons. *Speech communication, 39,* 33-46.

Effect of Personal Music Systems on Hearing in Young Adults

[1]Dhanalakshmi Ganesan & [2]Ajith Kumar Uppunda

Abstract

Personal music systems (PMS) such as iPods, MP3 players, portable compact disc (CD) players have become overwhelmingly popular among young people for well over a decade, and their popularity is increasing with advancements in technology. In developing countries like India, use of PMS, specifically, mobile phone MP3 usage is growing rapidly. The aim of the study was a) to measure the output levels of the PMS at the volume control setting that was preferred by the participant in quiet and in the presence of 65 dB SPL bus noise; b) to compare the extended high frequency hearing thresholds (3 kHz-20 kHz) in individuals who use PMS and individuals who don't; c) to compare the transient evoked otoacoustic emissions in individuals who use PMS and individuals who don't; d) to measure gap detection in noise, duration discrimination and modulation detection thresholds in individuals who use PMS and individuals who don't; e) to measure speech perception in noise in individuals who use PMS and individuals who don't. More than 90% of the participants in the current study used their PMS at levels exceeding the safety limits specified by Ministry of Environment and Forests. Hearing thresholds in extended high frequency, TEOAE amplitudes, and temporal processing and speech perception skills were significantly affected in PMS group than in control group. Poor performance of the PMS-group on temporal and speech perception measures indicate listening to loud levels of music may compromise central auditory system.

Keywords: *Personal music system, transient evoked otoacoustic emission, temporal processing tasks, speech perception tasks, output SPL.*

Introduction

Personal music systems (PMS) such as iPods, MP3 players, portable compact disc (CD) players have been very popular among young people for well over a decade, and their popularity is increasing with advancements in technology. While the use of these devices has become a main source of recreation for the young generation, the harmful effect of this habit on hearing ability has become an increasing concern in public health. A recent survey of large sample of adolescents and young adults revealed that many of them are developing hearing impairment at a young age (Chung, Des Roches, Meunier & Eavey 2005). One of the reason for this may be prolonged use of PMSs, which are typically played at a "loud" volume setting (Zogby, 2006). Prolonged exposure to high levels of noise/music can cause damage to the hair cells in the cochlea and results in permanent/temporary cochlear hearing loss.

When calculating the level of risk or amount of exposure, both duration of exposure and intensity of the signal must be considered. The National Institute for Occupational Safety and Health (1998) guidelines for work place settings specify that any exposure of 85 dBA for more than 8 hours exceeds the maximum daily allowable noise dose. As the intensity of the signal increases, the maximum allowable exposure duration decreases. While this standard is based on industrial noise, it is also currently used as the guideline for recreational noise exposure, including listening to music. In India, the Ministry of Environment and Forests (2000) has proposed a maximum allowable noise dose of 85 dBA for an 8 hour period per day. The Ministry of Environment and Forests (2000) recognizes that there is a tradeoff between the exposure time and the sound level, which is quantified by a '5 dB exchange rule'. Every 5 dB increase in the exposure level will be compensated by halving the exposure time to keep the risk constant. The maximum permissible level is also not fully harmless, as a few percent of people may still incur a permanent hearing loss if exposed to it. Typically, when discussing music listening behaviors, the general practice is to consider that noise levels exceeding the maximum daily dose as indicative of at-risk listening behavior (Fligor & Cox, 2004).

Recently, research has begun to focus on the risk of hearing loss during leisure-time or recreational activities, such as listening to music through PMSs (Biassoni, Serra & Richtert,2005; Chung et al., 2005). It has been shown in previous studies that the use of PMSs had damaging effects on hearing (Peng et al., (2007). Additionally, the newer style of headphone that accompanies the PMSs, known as 'ear bud', increases the problem by directly channelling the sound into the ear (Fligor & Cox, 2004). With the newer technology PMSs have the potential to be used for longer durations and at higher volumes than older technologies, as digital PMSs have expanded music storage and battery life capabilities and produce very low distortions even at high levels. The prolonged exposure to high levels of music can cause damage to the hair cells in the cochlea and results in permanent noise-induced cochlear hearing loss.

Otoacoustic emissions (OAEs) and hearing thresholds are highly susceptible to cochlear trauma, such as ex-

[1]Email: laxmiaiish19@gmail.com,
[2]Reader in Audiology, Email: ajithkumar18@gmail.com

posure to ototoxins or loud noises (Furst, Reshef & Attias, 1992). Kumar, Mathew, Alexander and Karan (2009) showed a significant correlation between increased pure-tone thresholds at 6 kHz and estimated 8-hour equivalent exposure in a group of young adults who were using PMSs. They also reported decreased amplitude of distortion product otoacoustic emission in the high frequencies in individuals who used PMSs when compared to those who did not. Recently, it has been shown that hearing thresholds at frequencies above 8000 Hz are more sensitive to noise exposure than the conventional audiometric frequencies (250 Hz - 8000 Hz). Testing frequencies higher than 8000 Hz may reveal the subclinical damage to the cochlear structures (Peng et al., 2007).

Psychophysical evidence indicates that presence of cochlear hearing loss causes deficits in temporal processing (Moore, 2007). Temporal processing encompasses a wide range of auditory skills including temporal resolution or temporal discrimination (i.e., gap detection and fusion). Normal perception of the temporal aspects of the stimulus is crucial for understanding speech in quiet and in adverse listening conditions. One of the important factors that contribute to the poor performance of hearing-impaired listeners on temporal processing tasks is audibility of the high-frequency signals.Apart from audibility, the suprathreshold distortions can also contribute to the poor performance of hearing-impaired listeners on temporal processing tasks. These suprathreshold distortions may be caused by changes in the central auditory system secondary to cochlear damage.

The importance of temporal resolution in speech perception has been studied by Drullman, Joost, Festen and Plomp (1994) in normal hearing individuals and found that temporal modulations <2 kHz were essential for accurate speech perception. Several studies have suggested that word and sentence recognition performances are both positively correlated with temporal resolution in hearing-impaired subjects especially in the presence of background noise (Eg: Glasberg & Moore, 1989).

In developing countries like India, use of PMS, specifically, mobile phone MP3 usage is growing rapidly. About six million new mobile subscriptions are added every month and three quarters of India's population was covered by a mobile network at the end of 2008 (Murthy, 2008). Output levels of the PMS can be as high as 121 dBA (Fligor & Cox, 2004). Kumar et al. (2009) reported significant positive correlation between hearing thresholds and exposed music levels at 6000 Hz suggesting that individuals who listened to music at high levels tend to have higher pure tone hearing thresholds at 6000 Hz. They also reported a significant negative correlation between high frequencies otoacoustic emission amplitudes and output levels of PMS suggesting that individuals who listened to music at higher levels had reduced OAE amplitudes. These relationships were observed even though all the participants in their study had "clinically normal" hearing thresholds and OAE amplitudes. These results suggest that, listening to music through PMS at higher intensities may cause subtle/pre-clinical damage to the outer hair cells (OHCs) and over the years such behavior may be hazardous to hearing. Subtle and subclinical damage to OHCs may lead to higher order perceptual problems.

Hence, this study was taken up to evaluate the effect of PMSs on peripheral auditory system by measuring the OAEs, extended high frequency hearing thresholds and the central auditory processing by evaluating the temporal processing tasks and speech perception in noise in clinically normal hearing individuals who regularly use PMS for more than 2 hours per day for 2 years or more. Specific temporal processing tasks that were evaluated included gap detection in noise, temporal modulation detection, and duration discrimination. Furthermore, this study also measured the output levels of the personal music systems at the level of the ear drum of the participants and compared that to damage risk criteria provided by the Ministry of Environment and Forests (2000).

Method

Participants

A total of 49 participants participated in the present research. They were divided into two groups based on their music listening habits. Group I consisted of 29 participants aged between 17 and 25 years (10 males, 19 females) who reported regular usage of personal music systems (PMS). Group II consisted of 20 age matched individuals who rarely ever listened to music through PMS. Hereafter, Group I will be called PMS-group and Group II will be called No-PMS group for easy nomenclature. Participants in both group had their air conduction and bone conduction hearing thresholds within 15 dB HL at octave frequency from 250 Hz to 8 kHz. All participants showed 'A' type tympanogram with acoustics reflex at normal sensation levels. None of them reported any history of middle ear pathology, ototoxic drugs usage or exposure to occupational noise.

Procedure

The study was conducted in 2 phases.

Phase 1: Measurement of output sound pressure levels (SPL) of PMS

Only PMS-group participated in this phase of the study. Output SPLs produced by the PMS were measured in the subject's ear canal using a probe microphone. A commercially available real ear probe microphone measurement system (Fonix, 7000) was used for this purpose. Insertion depth of the probe was 28 mm from the tip of the tube to tragal notch. This insertion depth is the standard insertion depth used while doing real ear probe microphone measurements in adults. All the measurements were made with participants' own PMS and a standard ear bud type of the earphones. After placing the probe tube in the ear canal, the earphones were placed. Participants were asked to play one of their favorite songs. Output SPLs were measured in two different conditions. In, first condition, participants were asked to adjust the volume control to their preferred listening setting in quiet. In the second condition the same was done in the presence of 65 dB SPL bus noise. Bus noise was considered as background noise as this condition is more naturalistic since most of the participants reported that they listen to music while commuting. Level of the noise produced by the bus engine in normal city ride condition (third gear at a speed of 40 kilometres per hour) at 2 feet from the driver (corresponds to 2-3 row of seat) was 65 dB SPL. Hence this condition was used to measure output SPL of the PMS.

Position of the probe microphone was not changed between any of the measurements. Output SPLs were measured at different frequencies from 200 Hz to 8000 Hz (over all 64 frequencies). These ear canal SPLs were converted to equivalent diffuse field SPLs to which the ear was exposed, by subtracting the transfer function of the open ear. This transformation is required to compare the output of PMS to damage risk criteria. As there is no evidence based definition for hazardous sound levels of music, as a substitute, standards for exposure to occupational noise have been proposed for use. The occupational noise exposure defines the maximum allowable noise levels in terms of diffuse field values and not as ear canal sound pressure levels. Hence, the ear canal sound pressure levels were converted into to diffuse field levels by subtracting the transfer function of the open ear canal. The transfer function of the open ear was obtained by calculating the difference between the reference location at the opening of the ear canal and the probe microphone SPL near the eardrum for a sweep frequency tone presented at 65 dB SPL. The output SPLs at individual frequencies was converted to dBA values by adding the A-weighting adjustment values. The overall SPL in dBA was then determined by logarithmically adding dBA values at each frequency. From this data 8 hour equivalent continuous A-weighted noise exposure (Leq8h) was calculated using the following equation:

$$Leq8h = \mathrm{L}_T + 10 \log_{10} [\mathrm{T}/8]$$

Where Leq8h is the 8 hour equivalent continuous noise exposure, T is the exposure time in hours (music listening hours per day); L_T is the Level of noise exposure over the time period T. The procedure followed for calculation of Leq8h from the ear canal sound pressure level was same as that described in Kumar, Mathew, Alexander and Kiran (2009).

Phase II: Auditory Measures

Extended high frequency audiometry, otoacoustic emissions, temporal processing tests and speech perception in noise were assessed in phase II. Both the PMS group and No-PMS group participated in these experiments.

Extended High Frequency Audiometry

Calibrated two channel diagnostic audiometer GSI 61 with transducer was HDA 200 was used for extended high-frequency audiometry. Using modified version of Hughson and Westlake procedure (Carhart & Jerger, 1959) pure tone hearing thresholds were estimated at different frequencies from 3 kHz to 20 kHz.

Otoacoustic Emissions

Transient evoked otoacoustic emissions (TEOAEs) were recorded using commercially available otoacoustic emission analyzer (ILO-V6). Participants were asked to sit on a reclining chair. TEOAE probe was inserted into their ear canal and TEOAEs were measured for 80 dB peak SPL clicks. Average response from a total of 260 clicks was used for the analysis. The overall TEOAE amplitudes and amplitudes at different frequency bands were noted and used for analysis.

Temporal Processing Tasks

Stimuli: All temporal processing tests except for the duration pattern test were carried out using mlp (maximum likelihood procedure) toolbox, which implements mlp in Matlab (Grassi & Soranzo, 2008).

General Procedure: The maximum likelihood procedure employs a large number of candidate psychometric functions and after each trial calculates the probability (or likelihood) of obtaining the listener's response to all of the stimuli that have been presented given each psychometric function. The psychometric function yielding the highest probability is used to determine the stimulus to be presented on the next trial. Within about 12 trials, the maximum likelihood procedure usually converges on a reasonably stable estimate of the most likely psychometric function, which then can be used to estimate thresholds. Stimuli were recorded at 44,100 Hz sampling rate. A two-interval alternate forced-choice method using a maximum length procedure (mlp) was employed to track an 80% correct response criterion. Thirty trials were used in the present study. During each trial, stimuli were presented in each of two intervals: One interval contained a reference stimulus, the other interval the variable stimulus. The participant indicated after each trial which interval contained the variable stimulus. This procedure was used in all temporal processing tests except for the duration discrimination test.

Gap detection thresholds: The participant's ability to detect a temporal gap in the center of a 750 ms broadband noise was measured. The noise was 0.5 ms cosine ramps at the beginning and end of the gap. In a two-interval alternate forced-choice task, the standard stimulus was always a 750 ms broadband noise with no gap, whereas the variable stimulus contained the gap.

Duration discrimination test: In this procedure, the minimum difference in duration that is necessary to perceive the two otherwise identical white noise bursts was measured. Duration of the standard stimulus is 250 ms the duration of the variable stimulus was changed according to the subject's response. In the two-interval alternate forced-choice procedure, the participants were instructed to tell which interval contained the longer duration signal.

Modulation detection thresholds for sinusoidally amplitude-modulated noise: Temporal modulation refers to a recurring change (in frequency or amplitude) in a signal over time. A 500 ms Gaussian noise was sinusoidally amplitude modulated at modulation frequencies, 8 Hz, 20 Hz, and 60 Hz and at 200 Hz. Noises had two 10 ms raise cosine ramps at onset and offset. Participants had to detect the modulation and tell which interval had the modulated noise. Modulated and un-modulated stimuli were equated for total root mean square power. Depth of the modulated signal was varied according to the subject's response up to an 80% criterion level. The modulation detection thresholds were expressed in dB by using the following relationship:

Modulation detection thresholds for sinusoidally amplitude-modulated noise: Temporal modulation refers to a recurring change (in frequency or amplitude) in a signal over time. A 500 ms Gaussian noise was sinusoidally amplitude modulated at modulation frequencies, 8 Hz, 20 Hz, and 60 Hz and at 200 Hz. Noises had two 10 ms raise cosine ramps at onset and offset. Participants had to detect the modulation and tell which interval had the modulated noise. Modulated and un-modulated stimuli were equated for total root mean square power. Depth of the modulated signal was varied according to the subject's response up to an 80% criterion level. The modulation detection thresholds were expressed in dB by using the following relationship:

$$\text{Modulation detection thresholds in dB} = 20log_{10}(\text{m})$$

Where m=modulation detection threshold in percentage.

Speech Perception in Noise

Speech perception in noise was evaluated using the speech in noise test developed by Avinash, Methi and Kumar, (2010). Seven equivalent lists from the original test were selected for the present study. Each list contained 7 sentences mixed with the eight talker speech babble noise at different signal to noise ratios (SNRs). First sentence in the each list was at +20 dB SNR, second sentence was at +15 dB SNR, third sentence was at +10 dB SNR, fourth sentence was at +5 dB SNR, fifth sentence was at 0 dB SNR, sixth sentence was at -5 dB SNR and last sentence was at -10 dB SNR. Each sentence had 5 key words. These sentences were presented through a personal computer (Lenovo Z372) at comfortable levels using a commercially available headphone (Index Mega supra-aural headphone, HS-301B). The listener's task was to repeat the sentences presented and each correctly repeated key word was awarded one point for a total possible score of 35 points per list.

Statistical Analysis

Appropriate statistical analysis was carried out on the data obtained to verify the objectives of the study.

Results

Measurement of Output Sound Pressure Levels (SPL) of Personal Music Systems (PMS)

Figure 1 shows mean output of the PMS at the entrance of ear canal (sound pressure level picked up by the probe microphone - head related transfer function) produced by the PMS at preferred listening levels across different frequencies (200 Hz - 8k Hz) in quiet and in the presence of 65 dB SPL bus noise. As can be seen from the Figure 1; in the presence of 65 dB SPL bus noise, preferred SPLs were higher when compared to quiet condition. Preliminary inspection of the data

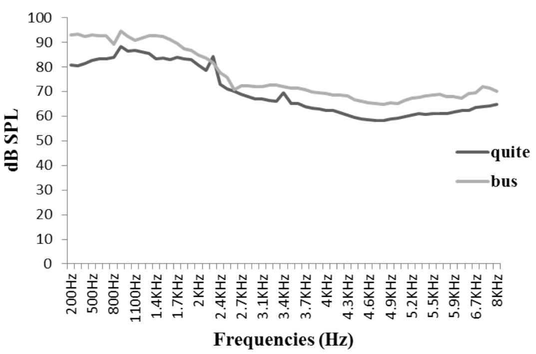


Figure 1: Mean output SPL level used by the participants at preferred listening setting in quiet, in the presence of 65dBSPL bus noise. Mean output in dB SPL is depicted at different frequencies. Please note that values in this figure are not converted into dBA and 8 hours equivalent level exposure.

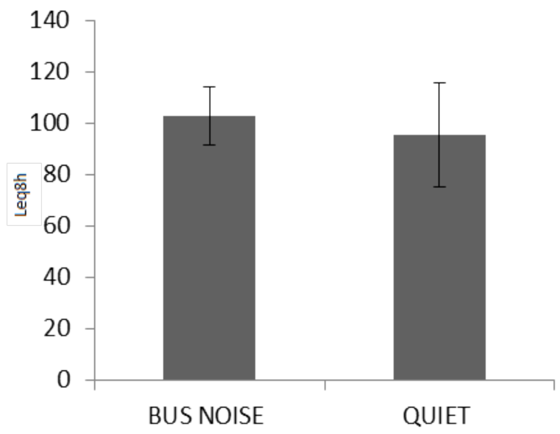


Figure 2: Mean Leq8h in quiet and in the presence of 65dB SPL bus noise. Error bars show 1 standard deviation.

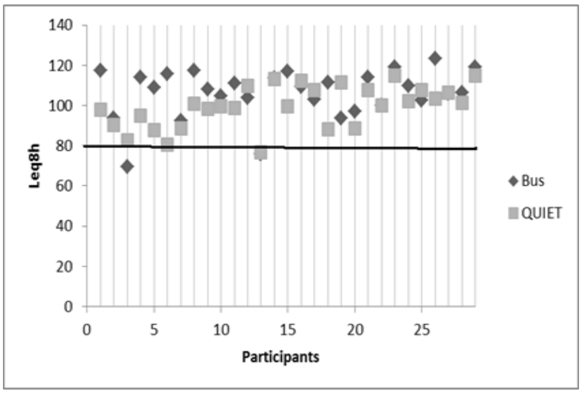


Figure 3: Output Leq8h values at participants preferred volume control setting in quiet and in the presence of bus noise for individual subjects. A square indicate the Leq8h in quiet and diamond indicate the Leq8h in the presence of bus noise.

shows that at low frequencies, output SPLs at preferred listening levels were higher than 80 dB SPL.

Figure 2 shows the mean Leq8h (8 hours equivalent level exposure) in quiet condition and in the presence of 65 dB SPL bus noise along with one standard deviation of variation. As can be seen from the Figure 2, mean 8 hour equivalent listening levels in quiet was 95.8 dBA and in presence of bus noise was 103.43 dBA. A paired 't'-test performed to see the significance of difference between mean Leq8h SPLs between two conditions. Results showed significant difference in mean Leq8h between two conditions ($t=3.062$, $p<0.05$). Figure 3 shows the Leq8h for individual participants. Squares indicate the Leq8h in quiet and diamond indicate the Leq8h in the presence of bus noise.

From the Figure 3, it can be inferred that most of the individuals' preferred listening levels increased in the presence of bus noise (compare the square and diamond Leq8h for each participant). Preferred listening Leq8h ranged from 63.3 to 114.8 dBA in quiet condition and was in the range of 72.8 to 122.1 dBA in the presence of 65 dB SPL bus noise. Furthermore, from the Figure 3, it can be observed that 26 of 29 individuals listened to music through their PMS at levels higher than 80 dBA in quiet condition and 27 individuals in the presence of bus noise.

Auditory Measures

High frequency audiometry

Figures 4 and 5 show the mean hearing thresholds in PMS group and No-PMS group along with one standard deviation of variation for right and left ear respectively. From these figures, it can be inferred that mean hearing thresholds were higher in the PMS group compared to No-PMS group.

ANOVA was done to see the significance of difference between the hearing thresholds between the two groups. Results showed a significant main effect of participant group on hearing thresholds in both right [F (12, 36) =

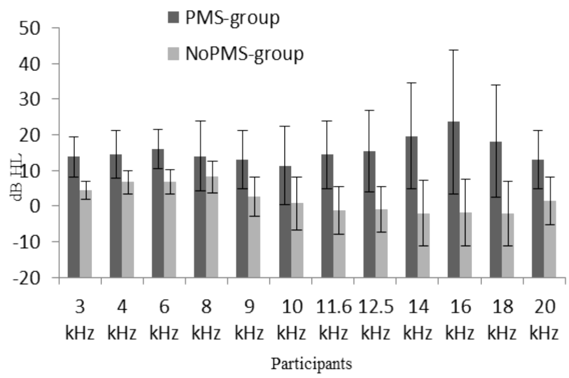


Figure 4: The mean high frequency hearing threshold across different frequencies in right ear for PMS group and No-PMS group. Error bar shows 1 standard deviation error.

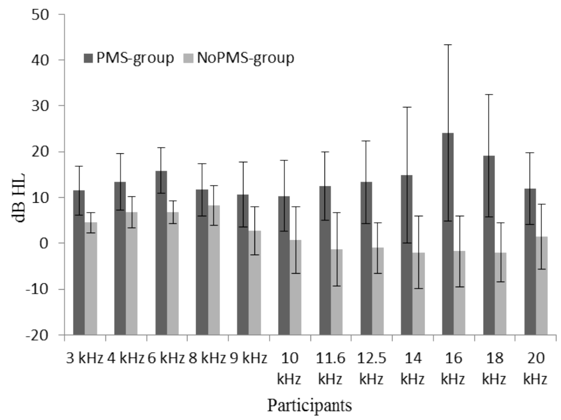


Figure 5: The mean high frequency hearing threshold across different frequencies in left ear for PMS users and NO-PMSs users. Error bars show 1 standard deviation error.

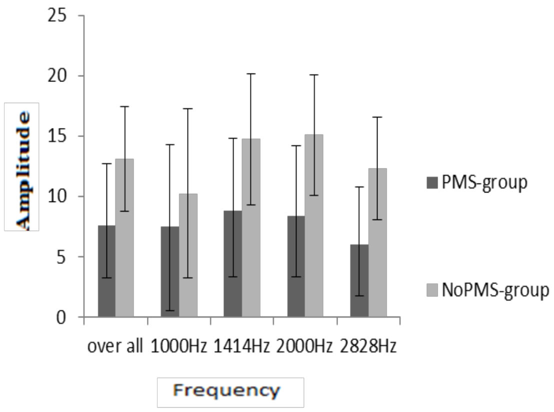


Figure 6: TEOAE amplitude across different frequencies in the right ear for PMS group and No-PMS group. Error bars show 1 standard deviation error.

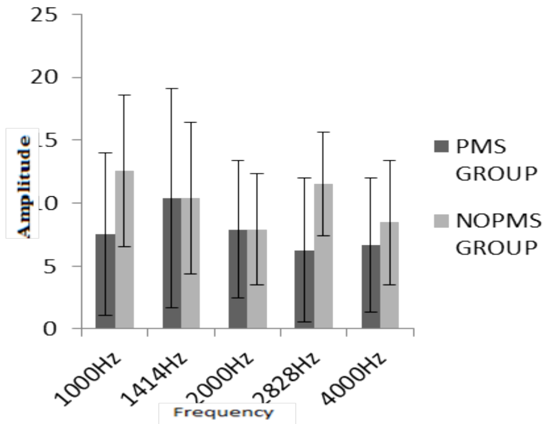


Figure 7: TEOAE amplitude across different frequencies in the left ear for PMS group and No-PMS group. Error bars show 1 standard deviation error.

Table 1: Significant difference between PMS group and No-PMS group in TEOAE results

Overall (SPL)	Right ear		Left ear	
	F	Significance levels	F	Significance levels
1000 Hz	4.23	0.045	9.48	0.003
1414 Hz	5.28	0.026	11.48	0.001
2000 Hz	2.61	0.05	0.002	0.107
2828Hz	148	0.001	11.42	0.001
4000 Hz	.298	0.11	3.70	0.05

8.5, $p<0.01$] and left ear [$F(12, 36) = 2.89$, $p< 0.01$]. Following pair wise comparison showed that at all the tested frequencies mean hearing thresholds of the PMS-group was significantly poorer than the mean hearing thresholds of No-PMS group.

Transient evoked otoacoustic emissions (TEOAEs)

Figures 6 and 7 show mean TEOAE amplitudes in two groups in right and left ear along with one standard deviation of error. Both the overall and band wise TEOAE amplitudes are shown in Figures 6 and 7. From these figures, it can be seen that both overall and band wise TEOAE amplitudes were reduced in PMS-group compared to No-PMS group. ANOVA was done to find the significance of difference between the mean TEOAE amplitudes between two groups. Results revealed that both the overall and band wise TEOAE amplitudes were significantly reduced in PMS group compared to No-PMS group in both right and left ear (except for 2 kHz in left and 4 kHz in right). F values and significance levels are depicted in Table 1.

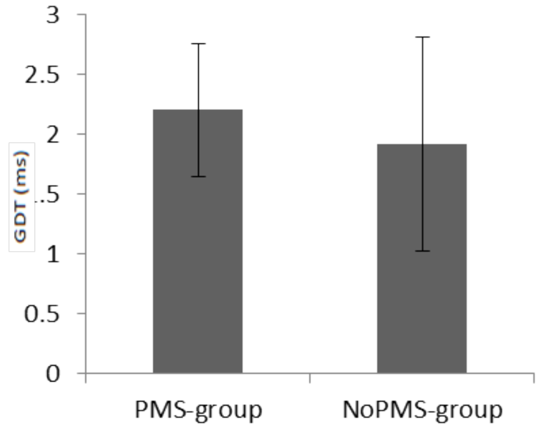


Figure 8: The mean gap detection threshold for PMS and No-PMS groups. Error bars show 1 standard deviation error.

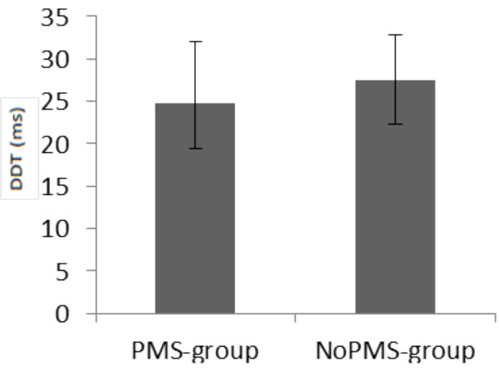


Figure 9: The mean duration discrimination threshold for PMS and No-PMS groups. Error bars show 1 standard deviation error.

Temporal processing

Figures 8 and 9 show the mean gap detection thresholds and duration discrimination thresholds in both the groups along with one standard deviation of variation. Independent samples 't' test was performed to see the significance of difference between mean gap detection thresholds and duration discrimination thresholds between two groups. Results showed a significant difference in mean gap detection thresholds between two the groups ($t=2.09$, $p<0.05$). However, there was no significant difference between the two groups in terms of du-

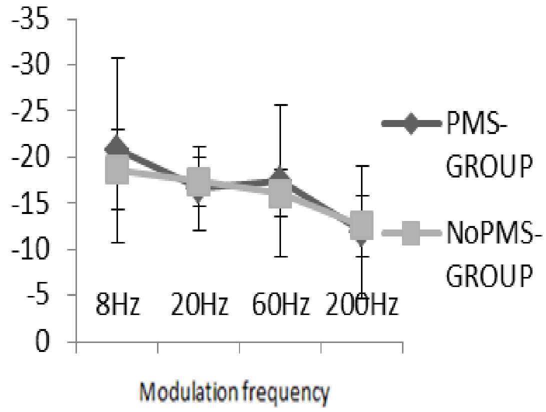


Figure 10: The mean sinusoidally amplitude modulation threshold for across different modulation frequency for PMS and No-PMS groups. Error bars show 1 standard deviation error.

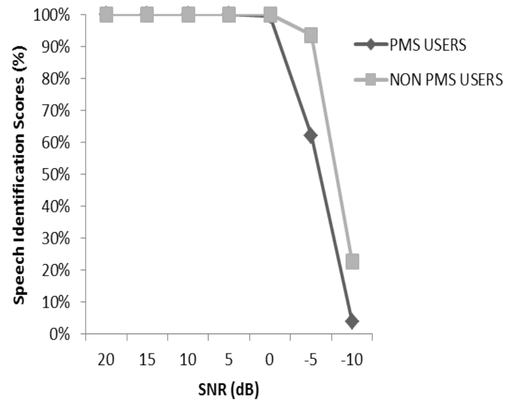


Figure 11: The mean percentage for PMS group and No-PMS group across different SNR levels.

ration discrimination thresholds (t=0.43, $p > 0.05$). Figure 10 shows the average modulation detection thresholds along with the one standard deviation of variation in both the groups. ANOVA did not reveal a statistically significant main effect of participant group on modulation detection threshold at 8 Hz [F (1, 33) =1.9, $p > 0.05$], 20 Hz [F (1, 33) = 0.06, $p > 0.05$] and 60 Hz. [F (1, 33) =1.1, $p > 0.05$]. However, at 200 Hz modulation frequency there was a significant main effect of participant group [F (1, 33) = 6.14, $p< 0.05$] on modulation detection thresholds.

Speech perception in noise

Figure 11 shows the mean scores for speech perception in noise (SIN) along with the one standard deviation variation for the both the groups. The mean scores indicate that the speech perception in noise is better for the No-PMS group as compared to the PMS group especially at higher signal to noise ratios (SNRs). The raw speech perception scores were converted in rationalized arcsine units (rau). All the further statistical analysis was carried out using the rau speech perception scores. At SNRs, +20 dB, +15dB, +10 dB, +5 dB and 0 dB, all participants in both the groups obtained 100% correct identification and hence these SNRs were excluded from further statistical analysis. To see the significance of differences in the speech perception scores between two groups, ANOVA was performed with speech scores at -5 dB and -10 dB SNR as dependent variable, taking participant groups as independent variable. Results showed a significant main effect of group on speech identification scores at -5 dB SNR [F (1, 47) = 6.997, $p<0.05$], and -10 dB SNR [F (1, 47) = 6.09, $p<0.05$].

Relationship between Leq8h and hearing measures

To evaluate the relationship between output SPLs and auditory measures, Pearson's Product-Moment correlation was carried out between participants' pure tone thresholds, temporal processing measures and speech perception in noise scores as dependent variable and Leq8h measured in the presence of bus noise as the independent variable. Results showed that none of the hearing measures had significant relationship with the leq8h.

Discussion

The purpose of this study was to measure the output levels of the PMS at the volume control setting that was preferred by the participants in quiet and in the presence of bus noise. Furthermore, this study also evaluated the extended high frequency hearing thresholds, temporal processing and speech perception skills in group of individuals who used PMS and compared that to individuals who don't use PMS. The mean dBA Leq8h at preferred volume control settings was 95.8 dBA and ranged between from 63.3 to 114.8 dBA in quiet. These preferred listening levels are slightly higher than what participants selected as " sounds best to you" in the Hodgetts, Rieger and Szarko (2007), or " medium/comfortable" in the Torre (2008). These numbers are alarming as more than 90% of the young adults who uses PMSs listens to music at sufficiently higher levels that can cause temporary/permanent hearing loss.

In the presence of bus noise of 65 dB SPL the mean preferred listening levels were increased to 103.43 dBA (range 72.8 to 122.1 dBA). This increase in the output levels in the presence of background noise is comparable to Hodgetts et al. (2007) who reported that participants increased the level of the music approximately 6 to 10 dB when either street noise or multitalker babble was added to the listening environment. The bus noise condition used in the current study is more naturalistic as most of the participants reported that they listen to music while commuting. No evidence based definition exists for hazardous sound levels of music. As a substitute, standards for exposure to occupational noise have been proposed for use.

In India, the Ministry of Environment and Forests (2000) has proposed a time weighted average level of 80 dBA for an 8 hour period per day as the maximum permissible limit. '5 dB exchange rule' has been proposed by the Ministry of Environment and Forests as a trade-off between exposure time and sound level. Referencing this criterion about 93.1% individuals in a group of 29 young adults listened to music above the safety limits prescribed by Ministry of Environment and Forests (2000). This is an alarmingly large proportion as majority of the individuals who listen to music through PMS may be putting themselves at risk for permanent noise induced hearing loss if exposed for extended periods of time (years). The output SPLs measured in the current study is slightly higher than what is reported by Kumar et al. (2009). This discrepancy is possibly because of the differences in the type of the earphones used in both the studies. Kumar et al. (2009) used the participants own earphones which ranged from head phones to ear bud type of receivers. In the current study a standard ear bud type of receiver with the deep insertion which is widely used with the current day PMSs was employed. Fligor and Cox (2004) have reported that ear bud type of the receiver produces higher sound pressure level (SPLs) than the other types of the receivers.

Effect of PMSs on Auditory Measures

High Frequency hearing thresholds

Results of the high frequency pure tone audiometry showed that hearing thresholds of individuals who used PMSs were significantly poorer compared to individuals who did not use PMSs. It should be noted that all the individuals in the PMS group had hearing thresholds within normal limits in the conventional audiometric frequency range. Extended high frequency hearing thresholds are reported to be more sensitive to noise induced damages than the conventional audiometric frequencies. Peng et al. (2007) reported that extended high frequencies may be affected by the noise earlier when compared to conventional audiometric frequencies. Poorer thresholds in extended high frequency region in combination with the clinically normal hearing thresholds in the conventional audiometric range reveals that listening to music through PMS at higher intensities may cause subtle pre-clinical damage to the auditory system and over the years such behavior may be hazardous to hearing.

Otoacoustic emissions

Overall transient evoked otoacoustic emission amplitudes were significantly poorer in individuals who use PMSs compared to those who don't. Miller, Marshall, Heller and Hughes (2006) reported that amplitudes of distortion product otoacoustic emission (DPOAE) are more sensitive to noise induced hearing loss than pure tone hearing thresholds. Kumar et al. (2009) reported a negative correlation between DPOAE measures and output SPLs of PMSs at preferred volume control settings. They concluded that individuals who listened to music at higher levels had reduced DPOAE amplitudes and signal to noise ratios (SNRs) even though the DPOAE amplitudes were within the clinical norms.

Temporal processing and speech perception in noise

Results revealed that both the temporal processing and speech perception (at higher SNRs) skills were adversely affected in PMS group. Given the present data, the observed deterioration in the temporal and speech processing skills in individuals who use PMS, in the presence of normal hearing sensitivity is probably due to changes in the central auditory system that was caused due to prolonged exposure to loud music through PMS. It has been reported that long-term noise may have persistent effect on brain function and behaviour, even when the peripheral hearing sensitivity is within normal range. Persistent effects of long-term noise exposure on central auditory system were evaluated using auditory evoked potentials. Kujawa and Liberman (2009) assessed the performance in visuo-motor target tracking task and simultaneously recorded the mismatch negativity for /pa/ and /ka/ contrasts on healthy individuals who were exposed to high levels of occupational noise. All their subjects had hearing thresholds that were comparable to the control group. Results showed impaired syllable-discrimination in the left hemisphere of noise-exposed individuals in silence and increased N2b complex for the novel sounds. Furthermore, attention control and ability to focus on visuo-motor tasks were aberrant in noise-exposed group. These results suggest that long-term exposure to occupational noise effects both sound discrimination mechanism and attention control mechanism.

Brattico et al. (2005)measured the neural responses in normal hearing, noise-exposed, and non-exposed participants to speech and non-speech deviants. Brain electrical source modelling suggested that speech sound contrast was lateralized to left hemisphere in non-noise expose group but in right hemisphere in noise-exposed group. This group differences were not found for the non-speech deviants. These studies show that long-term noise can have a detrimental effect on the central auditory system. This detrimental effect has been observed even when the peripheral hearing sensitivity is intact. The observed deficits in the temporal and speech processing abilities in normal hearing, PMS users in this study may be due to compromised central auditory system in these individuals. However, we cannot totally exclude the deleterious effects of distorted cochlear input as a factor. Normal hearing sensitivity does not necessarily mean the normal functioning of the cochlea in PMS groups.

Evidences from the animal research suggest that cochlear functioning can be affected even in the presence of normal-hearing sensitivity. In the current study, the amplitudes of the TEOAEs were significantly reduced in PMS compared to the group that did not use PMS. Kujala and Brattico (2009) reported a rapid and irreversible degeneration of spiral ganglion cells by the noise exposure which resulted in the temporary threshold shifts. Even though the hair cells and hearing sensitivity recovered with progression of time, the neuronal loss still persisted. Effects of such neuronal losses on auditory and speech processing are detrimental. In general, a trend of reduced temporal processing skills especially gap detection thresholds and speech perception in noise was observed in individuals who use PMSs.

Modulation detection thresholds for the high modulation frequency were significantly poorer in PMS group compared with control group. In the auditory system, modulations are represented by phase locked neural discharges of the auditory nerve fibres to individual cycles of modulation frequency. Data from the animal research have shown that acoustic over exposure can cause acute loss of afferent nerve terminals and degeneration of cochlear nerve. This might cause disruption in the phase locking and synchronization in the discharge patterns of auditory nerve fibres causing poor modulation detection thresholds. Poor modulation detection thresholds for higher modulation frequencies suggest that individuals in the PMS-group had difficulty in perceiving rapid fluctuations in the stimulus. Any complex broadband signals such as speech can be decomposed by auditory filters into relatively slow variations in the amplitude over time called envelop and relatively rapid oscillations called temporal fine structure. Importance of slowly varying temporal envelope in speech perception is well documented (Moore, 2008). Recently, it has also been demonstrated that temporal fine structure plays a crucial role in hearing in the presence of background noise (Oxenham & Simonson, 2009). It is necessary to perceive the rapid oscillations to derive benefits of temporal fine structure cues. Difficulty in perceiving rapid amplitude fluctuations in PMS group may also pose problems in coding temporal fine structure. This may be one of the reasons for poor performance of PMS group in speech perception measures. It is also been suggested that speech is co-modulated at the rate of fundamental frequency (in this study mean fundamental frequency of the target stimulus was 211 Hz). It is important to perceive these co-modulations to perceptually separate target speech and background babble as different acoustic streams. Difficulty of noise-exposed individuals in perceiving the rapid amplitude fluctuations may limit their ability to perceptually segregate target and background babble.

To summarize, the current study demonstrated that majority of young adults who use PMS, listen to music at hazardously high levels. This behavior did not result in clinically significant auditory changes in the present group of individuals using PMS. However, the individuals demonstrated poor performance in temporal and speech perception skills and reduced OAE and elevated hearing thresholds. All the results in combination suggests that listening to music through PMS at high levels results in compromised functioning of the cochlea and cerebral auditory systems which may not be reversible.

Conclusion

From this study, it can be concluded that alarmingly large proportion of the individuals who listen to music through PMS may be putting themselves at risk for permanent noise induced hearing loss if exposed for extended periods of time (years). Results of extended high frequency audiometry and otoacoustic emission showed that listening to music through personal music system at preferred volume control settings may not result in "clinically significant" elevation of hearing thresholds or reduction of otoacoustic emission amplitudes but may cause subtle pre-clinical damage to the auditory system and over the years such behavior may be hazardous to hearing.

Poor performance of the PMS group on temporal processing and speech perception measures indicate that listening the loud levels of music may compromise central auditory system. It should also be noted that in this study output levels were measured in the ear canals of young adults. The output sound pressure levels produced by these personal music systems in children may be significantly more as they have smaller ear canals. Specific evidence-based and theory-based studies, preferably longitudinal studies, should be conducted in young adults who use personal music systems, to develop effective intervention.

References

Avinash, M. C., Methi, R. R., & Kumar, A. (2010). Development of sentences for quick speech-in noise (Quick Sin) test in kannada. *Journal of Indian Speech and Hearing Association, 24*(1), 59-60.

Brattico, E., Kujala, T., Tervaniemi, M., Alku, P., & Monitillo, A. V. (2005). Long term exposure to occupational noise alters the cortical organization of sound processing. *Clinical Neurophysiology, 116*(1), 190-203.

Biassoni, E. C., Serra, M. R., & Richtert, U. (2005). Recreational noise exposure and its effects on the hearing of adolescents. Part II: development of hearing disorders. *International Journal of Audiology, 44*, 74-85.

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

Chung, J.H., Des Roches, C.M., Meunier, J., & Eavey, R. D. (2005). Evaluation of noise-induced hearing loss in young people using a Web-based survey technique. *Pediatrics, 115,* 861-867.

Drullman, R., Joost, M., Festen., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *Journal of the Acoustical society of America, 95*, 1-5.

Fligor, B. J., & Cox, L. C. (2004). Output levels of commercially available portable compact disc players and the potential risk to hearing. *Ear and Hearing, 24,* 513-527.

Furst, M., Reshef, I., & Attias, J. (1992). Manifestations of intense noise stimulation on spontaneous otoacoustic emissions and threshold microstructure: Experimental model. *Journal of the Acoustical Society of America, 91,* 1003-1014.

Glasberg, B. R., & Moore, B. C. J. (1989). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian Audiology, 32,* 1-25.

Hodgetts, W. E., Rieger, J. M., & Szarko, R. A. (2007). The effects of listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player. *Ear and Hearing, 28*(3), 290-297.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after temporary noise-induced hearing loss. *Journal of Neuroscience, 29*(45), 1477-85.

Kujala, T., & Brattico, E. (2009). Detrimental noise effects on brain's speech functions. *Biological Psychology , 81*(3), 135-43.

Kumar, A., Mathew, Alexander, S. N., & Kiran. (2009).Â Volume output sound pressure levels of personal music systems and their effect on hearing. *Noise and Health, 11*(44), 132-140.

Kumar, A., Ameenudin, S., & Sangamanatha, A.V.(2012). Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. *Noise and Health, 14*(58), 100-105.

Miller, J. A., Marshall, L., Heller, L. M., & Hughes, L. M. (2006). Low-level otoacoustic emissions may predict susceptibility to noise-induced hearing loss *Journal of the Acoustical Society of America, 120,* 280-96.

Ministry of Environment and Forest. (2000) S. O. 123 (E), [14/2/2000] Noise pollution (Regulation control) Rules. 2000. Available from www.envfor.nic.in/legis/legis.htm/#k.

Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological, and Technical Issues.* Chichester: John Wiley and Sons.

Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing impaired people. *Journal of the Association for Research in Otolaryngology, 9*(4), 399-406.

Murthy, R. (2008). Mobile based primary health care system for rural India. Retrieved from http://www.w3.org/2008/02/MS4D_WS/papers/cdac-mobile-healthcare-paper.pdf.

National Institute of Occupational Safety and Health (1998), NIOSH publication no, 98-126. http://'www.cdc.gov/nioshonus (mosn).

Oxenham, A. J., & Simonson, A. M. (2009). Masking release for low- and high pass- filtered speech in the presence of noise and single-talker interference. *Journal of the Acoustical Society of America, 125*(1), 457-468.

Peng, C. Y., Yajima, H., Burns, C. E., Zon, L. I., Sisodia, S. S., Pfaff, S. L.,et al. (2007). Risk of damage to hearing from personal listening devices in young adults. *The Journal of Otolaryngology, 36,* 181-185.

Torre, P. (2008). Young adults use and output level settings of personal music systems. *Ear and Hearing, 29,* 791-9.

Zogby (2006). Survey of teens and adults about the use of personal electronic devices and head phones. *American Speech-Language-Hearing Association*. http://www.asha.org/uploadedFiles/about/news /atitbtot/zogby_survey2006.pdf

Action Potential Latency in Individuals with Endolymphatic Hydrops

[1]Divya Vishu & [2]Niraj Kumar Singh

Abstract

Meniere's disease has one of the highest prevalence among those that affect the peripheral vestibular system. Various tests have been employed for the diagnosis of Endolymphatic hydrops. These include Glycerol test, OAE, VEMP, CHAMP, and ECochG in addition to the conventional routine audiological evaluations. However, the reports in the literature are suggestive of inaccurate and inadequate performance of these tests in the diagnosis of Meniere's disease. The present study aimed at checking the utility of latency difference between condensation and rarefaction polarities of AP and ABR in the diagnosis of Meniere's disease. The study also aimed at comparing these two techniques with the more established SP/AP amplitude ratio. In the present study, the conventional ABR and extratympanic ECochG were recorded from 21 ears of individuals with Meniere's disease with pure tone average less than 55 dB, 25 ears of individuals with sensorineural hearing loss other than sloping configuration and pure tone average less than 55 dB, and also 48 ears of healthy individuals. The latency of AP and wave I of ABR for rarefaction and condensation, and SP/AP ratio were measured for all the group of participants. There was a significant correlation for the latency difference between AP polarities and wave I of ABR. There was no correlation between SP/AP ratio and latency difference between AP polarities and also between SP/AP ratio and wave I of ABR. The AP latency difference and wave I latency difference produced higher positive results in Meniere's disease than the SP/AP ratio. A combination of the AP latency difference and SP/AP ratio could identify Meniere's disease in 85% of individuals and hence the combination would be a better choice.

Keywords: *ECochG, SP/AP amplitude ratio, Meniere's disease.*

Introduction

Electrococheography (ECochG) is a technique of recording stimulus related responses or the electrical potentials of the inner ear and auditory nerve. It is employed to evaluate cochlear function in patients with Meniere's disease. The underlying pathologic finding in Meniere's disease is widely suspected to be endolymphatic hydrops, which has been shown in animal studies to systematically alter cochlear potentials (Kimura, 1982; Aran, Rarey & Hawkins, 1984).

The cochlear potentials of interest in clinical ECochG are eighth nerve compound action potential (AP), summating potential (SP) and cochlear microphonics (CM). The AP results from simultaneous, stimulus-locked discharge of a population of spiral ganglion neurons (Kiang, Watanabe & Thomas, 1965; Cullen, Ellis & Berlin, 1972). The SP is a stimulus-locked direct current potential that can be observed as a baseline shift in the CM, and is also generated by cochlear hair cells (Dallos, 1973). The CM is an electrical response that mimics the acoustic waveform of the stimulus and is generated by the cochlear hair cells (Dallos, 1973).

A variety of electrode locations have been employed to record these potentials in animal and human investigations. In animal studies, electrodes are commonly placed in the cochlea (Van Deelen & Smoorenburg, 1986), on the round window (Prijs, 1985) or directly on the auditory nerve (Kiang, 1965). In humans, three electrode sites have been employed. Transtympanic ECochG is performed by placing a needle electrode through the tympanic membrane and onto the promontory (Moffat, Gibson, Ramsden, Morrison & Booth, 1977). Tympanic ECochG employs an electrode that is placed on the tympanic membrane (Margolis, Rieks, Fournier & Levine, 1995). Extratympanic ECochG is performed with an electrode placement in contact with the ear canal wall (Mori, Asai, & Matsunaga, 1987). These electrode sites tend to impact the morphology of the thus recorded ECochG waveform. In general, response amplitudes diminish with increasing distance from the cochlea (Eggermont, Odenthal, Schmidt, & Spoor, 1974).

ECochG has been used for the diagnosis of several auditory pathologies. These include auditory dys-synchrony (Roland, Yellin, Meyerhoff & Frank, 1995;Santarelli & Arslan, 2002; Anastasio, Alvarenga, & Filho, 2008) and also Meniere's disease (Aso, Watanabe & Mizukoshi, 1991; Mori, Asai, Suizu, Ohta & Matsunaga, 1985; Mori, Asai & Matsunaga, 1987; Ferraro & Tibbils, 1999; Saas, Densert, Magnusson & Whitaker 1998) among others.

The inadequacy of various tests in the diagnosis of Meniere's disease calls for further studies that could aid its diagnosis. So, the present study was conducted with the aim of evaluating the utility of latency difference between rarefaction and condensation polarities of AP and also ABR wave I in the diagnosis of Meniere's disease. The study also aimed at comparing the above techniques to identify the better of the three techniques in the diagnosis of Meniere's disease.

[1]Email: divyavishu27@gmail.com
[2]2Lecturer in audiology, Email: niraj6@gmail.com

Table 1: Protocol for recording ECochG and ABR

Stimulus parameter	ECochG	ABR	Acquisition parameter	ECochG	ABR
Stimulus intensity	90 dB nHL	90 dB nHL	Analysis time	5ms	10 ms
Stimulus rate	11.1/s	11.1/s	Pre-stimulus time	2ms	0 ms
Stimulus polarity	Rarefaction and condensation	Rarefaction and condensation	Amplification	50000 times	100000 times
Stimulus polarity	Clicks	Clicks	Filter settings	10 Hz -3000	100 Hz- 3000 Hz
			Sweeps	1000	1000

Method

The present study was conducted to evaluate the efficacy of latency difference of action potential between rarefaction and condensation polarities in the diagnosis of Meniere's disease. This was aimed at specifically using extratympanic recording technique of ECochG for the same.

Participants

The study incorporated three sets of participants who were divided into three groups; a Meniere's disease group, a sensorineural hearing loss group and a group of healthy individuals. Group I consisted of 21 ears of participants in the age range of 18 to 55 years (9 males & 12 females) who were diagnosed with endolymphatic hydrops based on the questionnaire of American Academy of Otolaryngology Head and Neck Surgery (1995). Each of the participants within this group had pure tone average threshold of less than 55 dB HL. Their unaffected ears served as a separate group for a number of analyses. The group II consisted of 25 ears of 16 participants (7 males & 9 females) with sensorineural hearing loss (other than Meniere's disease) in the same age range as group I. The other subject selection criteria for this group included exclusion of individuals with pure-tone average threshold exceeding 55 dBHL and sloping audiometric configuration. The existence of neural pathology was screened out using ABR. Forty eight ears of healthy individuals (age & gender matched to group I) with normal audio-vestibular system served as the participants in group III.

Instrumentation

A calibrated diagnostic audiometer GSI-61 with TDH-39 supra-aural headphones housed in MX-41/AR ear-cushions and Radioear B-71 bone vibrator was used for estimating air conduction and bone conduction thresholds. The same set of equipments in AC mode alone was used for speech audiometry.

A calibrated diagnostic immittance meter GSI-Tympstar was used to obtain Tympanogram. Same equipment was also used for obtaining ipsilateral and contralateral acoustic reflex thresholds (ARTs).

An Intelligent Hearing System Smart EP version 4.0 with ER-3A insert earphones connected with TIPtrode was used to acquire extratympanic ECochG. The same instrument without TIPtrode was used to acquire ABR.

Test Procedure

The routine audiological evaluation involved pure tone audiometry, speech audiometry, and immittance evaluation. Pure tone audiometry was done using the Carhart and Jerger (1959) modified Hughson and Westlake method for the octave frequencies of 250 through 8000 Hz for air conducted stimuli using TDH-39 headphones. Bone conduction thresholds were obtained for the octave frequencies of 250 through 4000 Hz. The word recognition score were obtained at the most comfortable level using the standardized word lists in the client's native language. Immittance evaluation was done to rule out any middle ear pathologies. It involved obtaining tympanogram and acoustic reflex thresholds (both ipsilateral and contralateral). Tympanograms were obtained using a 226 Hz probe tone frequency whereas the ARTs were obtained at frequencies from 500 Hz through 4000 Hz using the above mentioned probe tone frequency. The ABR was used to screen out neural pathology.

ECochG was administered by seating the participants comfortably in a well illuminated acoustically treated test room with the ambient noise levels within ANSI specifications (ANSI S3.1-1999). The skin overlying the electrode sites were cleaned using Nuprep skin preparing gel prior to the electrode placement. For the preparation of ear canal skin, the same skin preparing gel was used with a swab stick. The electrodes were mounted using Ten20 conduction gel and surgical plaster. The electrode montage consisted of TIPtrode as the non-inverting electrode which was placed in the ear canal; inverting electrode was placed on the test ear mastoid; and ground was placed on the forehead. The inverting and ground electrodes were the regular disc type silver chloride electrodes. An adult-size TIPtrode was attached to insertion cushion on the TIPtrode tubing. Tiptrode plug was then compressed tightly and placed in the ear canal while pulling the pinna upward, backward, and slightly outward, in a circular movement. It was ensured that the impedance for each electrode was less than 5kΩ and the inter-electrode impedance difference was less than 2 kΩ. The protocol for ECochG has been shown in table 1.

ABR was administered with the electrode montage that included the placement of inverting electrode on the test ear mastoid, non-inverting electrode on the forehead and ground on the non-test ear mastoid. All the electrodes were the regular disc type silver chloride electrodes. The participant preparation and the impedance values required for the electrodes for ABR were similar to that of ECochG. The protocol for ABR has been shown in Table 1.

Response Measure

The latency of the action potential and wave I of ABR for rarefaction and condensation, and SP/AP ratio were measured for all the group of participants. From that the shift in the latencies between the condensation and rarefaction polarities were measured by subtracting one from the other.

Statistical Analysis

A descriptive statistics was done to obtain the mean and standard deviation for the measures.Since the data obtained was non-normally distributed, the non-parametric statistical analysis was done. This involved a Kruskal Wallis test for overall comparison and a Mann-Whitney U test for pairwise comparison. A Kappa analysis was also done for checking the agreement between the SP/AP ratio and AP latency difference between the condensation and rarefaction polarities of click.

Results

The present study was conducted with the aim of checking the utility of latency difference between the condensation and rarefaction polarity of action potential in the diagnosis of Meniere's disease. In addition, it was aimed at evaluating the utility of a similar difference for ABR wave I in the diagnosis of Meniere's disease. Furthermore, the study also aimed at checking the efficacy of SP/AP ratio in the diagnosis of Meniere's disease and comparing this method to the AP latency difference to find out which is a better tool for the diagnosis of Meniere's disease. To fulfil these aims, the participants were divided in to 3 groups. The results are discussed under the headings of AP latency difference, SP/AP ratio, and ABR latency difference to compare between groups. Also a correlation between SP/AP ratio, ABR latency difference, and AP latency difference was evaluated. For several comparisons, the unaffected ears of individuals with Meniere's disease were considered as a separate group.

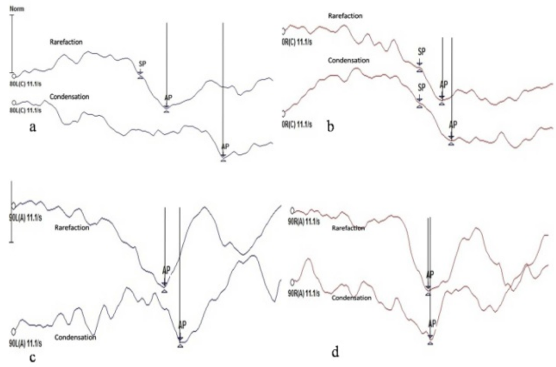


Figure 1: Representative AP waveforms (in alphabetic order) from affected ears with Meniere's disease, ears of healthy individuals, unaffected ears of individuals with Meniere's disease, and ears with sensorineural hearing loss.

AP Latency Difference between Condensation and Rarefaction Polarities of ECochG

All the participants within each of the three groups underwent ECochG and AP were identified in the condensation as well as the rarefaction polarities' waveforms. Sample ECochG waveforms from one participant from each of the groups are shown in Figure 1.

Statistical analysis was done using Statistical Prediction for the Social Sciences (SPSS) software version 17. The waveforms were analyzed for latencies of action potential for rarefaction and condensation polarities and the difference between the two was obtained. The values so obtained were then subjected to *descriptive analysis* to obtain mean and standard deviations. The *mean* and *standard deviation* values for ears of healthy individuals, ears of individuals with sensorineural hearing loss, unaffected ears of individuals with Meniere's disease and affected ears of individuals with Meniere's disease were found to be *0.13 ms (S.D = 0.02), 0.13 ms (S.D = 0.02), 0.21ms (S.D = 0.02), and 0.47ms (S.D = 0.11)* respectively. The mean of AP latency difference in the affected ears of individuals with Meniere's disease was higher than the other two groups and also compared to their own unaffected ears. Likewise, the unaffected ears of individuals with Meniere's disease also produced larger mean latency difference value than the ears of healthy individuals and also those of the individuals with sensorineural hearing loss. The same has been depicted in Figure 2.

A *Kruskal Wallis test* was administered to compare between the ears of healthy individuals, ears with SNHL, unaffected ears of individuals with Meniere's disease and affected ears of individuals with Meniere's disease in terms of the difference between the Action potential latencies between rarefaction and condensation clicks. The results revealed *significant difference* between the latencies of the two polarities [$\chi^2(3) = 71.889$, $p = 0.000$]. A *post hoc analysis* was done using *Mann-*

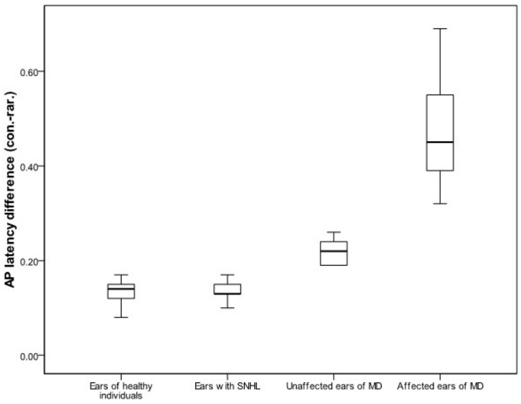


Figure 2: The box plot of AP latency difference between condensation and rarefaction polarities.

Whitney U test for pair wise comparison between all possible pairs. The pair wise comparison revealed a *significant difference* between all the pairs except between ears of healthy individuals and ears with SNHL. The latency difference between the polarities in the affected ears of individuals with Meniere's disease was also significantly different from the other groups. The latency difference was largest for the affected ears of individuals with Meniere's disease followed by their unaffected ears. The ears of healthy individuals and those of individuals with sensorineural hearing loss revealed lesser latency difference between the polarities than either of the above two and the two were comparable. The exact *'p'* and *'Z'* values are given in table 2.

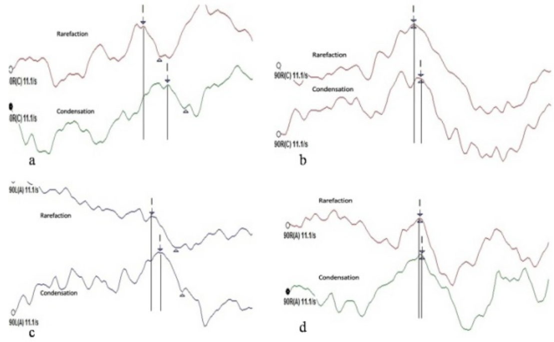


Figure 3: Panel 'a', 'b', 'c' and 'd' represent the ABR waveforms obtained from ears of individuals with Meniere's disease, ears of healthy individuals, unaffected ears of individuals with Meniere's disease, and ears with sensorineural hearing loss respectively.

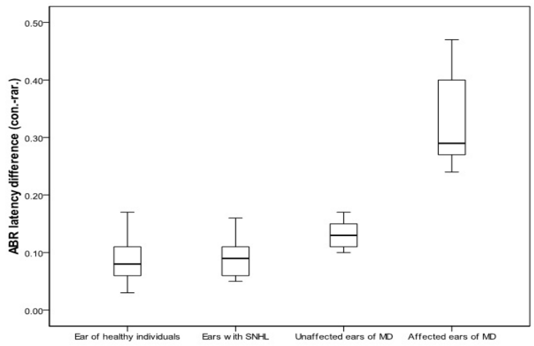


Figure 4: The box plot of the of ABR latency difference between condensation and rarefaction polarities.

ABR Wave I Latency Difference between Condensation and Rarefaction

All the participants within each of the groups underwent ABR and peaks (waves) were identified. Sample ABR waveforms from one participant from each of the groups are shown in Figure 3.

The waveforms were analyzed for latencies of wave I of ABR for rarefaction and condensation polarities and the difference between the two was obtained. The values thus obtained were then subjected to descriptive analysis. The ears of healthy individuals produced a *mean* ABR wave I latency difference of *0.08 ms (S.D. = 0.03)* between the two polarities used in the study. The difference for ears with sensorineural hearing loss, unaffected ears of Meniere's disease, and affected ears of Meniere's disease was *0.08 ms (S.D = 0.03), 0.13ms (S.D = 0.02),* and *0.32 ms (S.D = 0.07)* respectively. A graphical illustration of the same has been put forward in Figure 4.

A *Kruskal Wallis test* was administered to compare the groups in terms of the latency difference between rarefaction and condensation polarities for wave I of ABR. The results revealed *a significant difference* between the groups [$\chi^2(3) = 60.358, p = 0.000$]. A pair wise comparison was done using the *Mann-Whitney U test statistic* which revealed the latency difference of ABR be-

Table 2: 'Z' and 'p' values of Mann-Whitney U test for AP latency difference between condensation and rarefaction polarities

	Ears of healthy individuals	Ears of SNHL	Unaffected ears of MD	Affected ears of MD
Ears of SNHL	Z = -0.653	-	Z = -5.805	Z = -4.968
	p = 0.514	-	p = 0.000	p = 0.000
Unaffected ears of MD	Z = -5.530	-	-	Z = -4.843
	p = 0.000	-	-	p = 0.000
Affected ears of MD	Z = -6.601	-	-	-
	p = 0.000	-	-	-

Table 3: 'Z' and 'p' values of Mann- Whitney U test for ABR latency difference between condensation and rarefaction polarities

	Ears of healthy individuals	Ears of SNHL	Unaffected ears of MD	Affected ears of MD
Ears of SNHL	Z = -0.579 p = 0.563	- -	Z = -3.415 p = 0.001	Z = -5.798 p = 0.000
Unaffected ears of MD	Z = -3.749 p = 0.000	- -	- -	Z = -4.844 p = 0.000
Affected ears of MD	Z = -6.588 p = 0.000	- -	- -	- -

tween the polarities in the affected ears of individuals with Meniere's disease to be *significantly different* from all of the groups (ears of healthy individuals, ears with SNHL, & unaffected ears of individuals with Meniere's disease). The comparison between ears of healthy individuals and those of sensorineural hearing loss showed *no significant difference*. The latency difference between the polarities in the unaffected ears of individuals with Meniere's disease was *significantly different* from all others. The latency difference was greatest for the affected ears of Meniere's followed by their unaffected ears. The other two groups produced nearly equivalent latency differences. The 'Z' and 'p' values for the pairwise comparisons also have been shown in table 3.

SP/AP Ratio

The ECochG waveforms obtained from each individual were analyzed. The SP/AP amplitudes were obtained and their ratio was computed and subjected to *descriptive analysis*. The SP/AP amplitude ratio was highest for the affected ears of individuals with Meniere's disease [*Mean = 0.43, S.D = 0.19*]. The ears with sensorineural hearing loss [*Mean = 0.26, S.D = 0.07*] produced comparable SP/AP ratio values to the ears of healthy individuals [*Mean = 0.26, S.D = 0.07*] and also the unaffected ears of Meniere's disease [*Mean = 0.26, S.D = 0.07*]. A graphical representation of the same has been provided in Figure 5.

A *Kruskal Wallis test* was administered to compare the four groups in terms of SP/AP ratio. The results revealed a *significant difference* in SP/AP amplitude ratio between the groups [$\chi^2(3) = 11.31, p = 0.01$]. A *post hoc analysis* was done using *Mann-Whitney U test* for pair wise comparison between all possible pairs. The affected ears of Meniere's disease were found to be *significantly different* from all others on the pair wise comparison. This apart, there was *no significant difference* between other pairs. The *'p'* and *'Z'* values of pairwise comparison have been given in table 4.

Relationship between AP Latency Difference and SP/AP Ratio in the Affected Ears of the Individuals with Meniere's Disease

The present study aimed at evaluating a relationship between AP latency difference and SP/AP ratio in individuals with Meniere's disease. The *Spearman's correlation analysis* was used to obtain the relationship between the difference in AP latencies in rarefaction and condensation polarities and SP/AP ratio in the affected ear of MD. The results showed the existence of *slight negative correlation* (Viera & Garrett, 2005) between the two which was statistically *not significant* [r_s= -0.070, p = 0.797]. Figure 6 shows the scatter plot illustrating this relationship.

Another statistical tool used was the *Kappa coefficient analysis*. For this, the AP latency difference values were converted into categorical data of Meniere's and non-Meniere's disease ears by using the mean reference values from the available research regarding the two variables. The mean value of AP latency difference of ≥ 0.40ms (Orchik, Ge & Shea, 1998) and SP/AP ratio of ≥ 0.35 (Ohashi, Nishino, Arai, Hyodo & Takatsu, 2009) was used to categorize the data into Meniere's and Non-Meniere's ears. The SP was present in only 16 ears (out of 21 ears) of the individuals with Meniere's disease and hence *Kappa correlation analysis* was done using only these many ears. The results revealed *slight correlation* (Viera & Garrett, 2005) which was statistically

Table 4: 'Z' and 'p' values of Mann- Whitney U test of SP/AP ratio

	Ears of healthy individuals	Ears of SNHL	Unaffected ears of MD	Affected ears of MD
Ears of SNHL	Z = -6.14 p = 0.53	- -	Z = -0.45 p = 0.653	Z = - 2.314 p = 0.021
Unaffected ears of MD	Z = -0.77 p = 0.93	- -	- -	Z = -2.147 p = 0.032
Affected ears of MD	Z = - 2.991 p = 0.003	- -	- -	- -

not significant [*K = 0.127, p = 0.61*]. There was agreement for positive results of Meniere's disease for 7 ears (out of 16) and negative results for 2 ears. The overall agreement between tests for Meniere's disease diagnosis was only 56.25%. This implies that a correct diagnosis of Meniere's disease versus non-Meniere's disease was made in only 56.25% of individuals when using a positive criterion on both the methods. When the diagnosis of MD was based on the positive results on either of the two methods, the identification of MD increased to 85%.

Relationship between ABR Wave I Latency and SP/AP Ratio in the Affected Ears of Individuals with Meniere's Disease.

One of the objectives of the study was to check if the ABR wave I latency difference between condensation and rarefaction polarities could yield results that could help in the diagnosis of Meniere's disease. A *Spearman's correlation analysis* was used to correlate the difference in ABR wave I latencies and SP/AP ratio in the affected ear of individual with MD. The results showed *slight negative correlation* (Viera & Garrett, 2005) between the two which was statistically *not significant* [r_s*=-0.131,p = 0.630*]. The same has been illustrated in Figure 7.

Relationship between AP Latency Difference and ABR Wave I Latency Difference in the Affected Ears of Individuals with Meniere's Disease

A *Spearman's correlation analysis* was used to establish the relationship between the difference in AP latencies and ABR wave I latencies in the affected ears of individuals with Meniere's disease. The results revealed an *almost perfect positive correlation* (Viera & Garrett, 2005) between the two set of variables, and this was statistically *significant* [r_s*= 0.938, p = 0.000*]. The same has been illustrated in Figure 8.

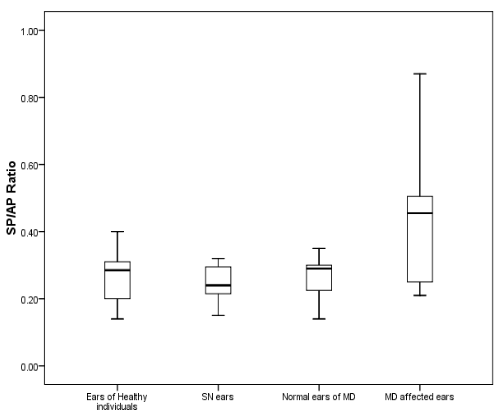


Figure 5: The box plot of the SP/AP amplitude ratio.

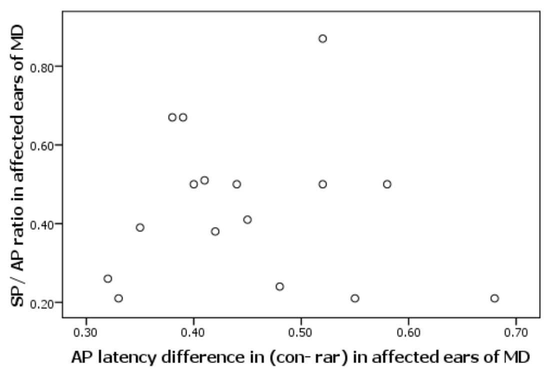


Figure 6: The scatter plot showing the relationship between SP/AP ratio and AP latency difference in the affected ears of individuals with Meniere's disease.

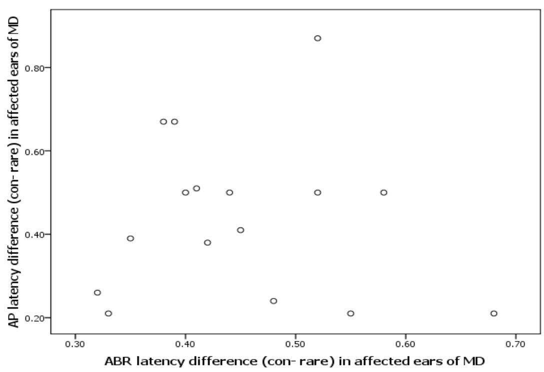


Figure 7: The scatter plot showing the correlation between SP/AP ratio and ABR wave I latency difference in the affected ears of individuals with Meniere's disease.

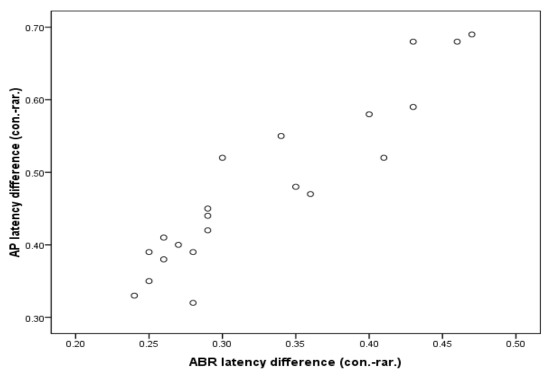


Figure 8: The scatter plot showing the relationship between AP latency difference and ABR wave I latency difference in the affected ears of individuals with Meniere's disease.

Discussion

AP Latency Difference between Condensation and Rarefaction Polarities of ECochG.

From the results of the present study it was evident that the mean AP latency difference between condensation and rarefaction polarities of AP was higher for the affected ears of the individuals with Meniere's disease compared to ears of healthy individuals, ears with sensorineural hearing loss and also their own unaffected ears. The present study did not compare the rarefaction and condensation polarities separately across the groups. However, the findings in literature have shown the difference to be prominent for condensation polarity (Saas, Densert, Magnussen, Whitaker, 1998). The authors reported no difference for rarefaction polarity between the healthy individuals, individuals with sensorineural hearing loss and those with Meniere's disease; however the individuals with Meniere's disease revealed longer latency of AP for condensation polarity than the other two groups of their study. This probably may be the reason for the larger difference between the latencies of the two polarities in individuals with Meniere's disease even in the present study. The prolongation of latency for condensation clicks alone may be explained on the basis of postulations of Tonndorf (1976) and the impact of hydrops on the travelling wave velocity (Eggermont & Odenthal, 1974). Tonndorf (1976), through his cochlear model, postulated that the basilar membrane, when loaded with endolymphatic hydrops, undergoes a downward displacement. Eggermont et al. (1974) studied the mode of the excitation in human cochlea and demonstrated that the latency of the AP was dependent on the velocity of the travelling wave, especially in cases with endolymphatic hydrops. The travelling wave was more affected in response to a condensation clicks than to rarefaction clicks due to the load of endolymphatic hydrops, and this resulted in a latency prolongation for condensation polarity. Therefore, the increased stiffness and change in position of the basilar membrane is likely to influence the pattern of the travelling wave motion, thereby aiding the explanation of the increased latency difference between the polarities in the individuals with Meniere's disease. The results of the present study are in agreement with the findings in the literature about the AP latency difference in Meniere's disease. Saas et al. (1998) conducted a study using transtympanic electrode placement on 30 individuals with Meniere's disease whose pure tone average thresholds ranged between 20 and 65dB. In addition, their study also used 11 patients with cochlear hearing loss of other aetiologies with pure tone average ranging between 30 and 60 dB and also 10 healthy subjects. They reported a mean difference in AP latencies between condensation and rarefaction of 0.02 ms, 0.01ms, and 0.32ms respectively in ears of healthy individuals, ears with sensorineural hearing loss, and ears of individuals with Meniere's disease. The findings of the present study revealed a similar pattern across the groups. However, the values obtained were higher (0.13 ms, 0.13 ms, and 0.47 ms in the same order as above). The values observed in the present study were appreciably higher for all the three groups of Saas et al. (1998). The differences in the values between the present study and Saas et al. (1998) could be related to the differences in the site of electrode placement. They used a transtympanic (near field) placement of the non-inverting electrode as against the extratympanic (far field) placement in the present study. Another study by Chen, Kang, Yeh and Wang (2004) used extratympanic electrode placement on ears of 10 healthy individuals and 33 individuals with Meniere's disease. They obtained a mean AP latency shift of 0.55 ms in the ears of individuals Meniere's disease and 0.11 ms in the ears of healthy individuals. The values observed in the present study were similar to those reported by Chen et al. which probably may be related to the use of similar method, including the extratympanic electrode placement in both the studies. So it could be concluded that the latency difference of AP between the polarities is greater in individuals with Meniere's disease irrespective of the site of electrode placement generally used for ECochG recordings. However, the cut-off values for Meniere's disease diagnosis could vary depending on the site of placement of electrodes. This calls for obtaining separate clinical values for different electrode placements for the diagnosis of Meniere's disease using ECochG.

The study revealed another finding of interest. The unaffected ears of individuals with Meniere's disease showed significant deviation in terms of latency difference between the polarities from the ears with sensorineural hearing loss and those with healthy individuals. The latency difference was*significantly* larger for the unaffected ears of individuals with Meniere's disease than their healthy and sensorineural counterparts. The findings in the literature has shown that, though Meniere's disease generally begins as a unilateral condition, it has a tendency to progress to the other ear within 2-7 years of its onset in the first ear in more than 50% of the individuals (Morrison, 1981., Salvinelli et al., 1998; Jackson & Silverstein, 2002., Saeed & Penny, 2011). Similar views were echoed by Huffelen, Mateijsen and Wit (1998) who obtained smaller OAE amplitudes in unaffected ears of individuals with Meniere's disease than the ears of normal hearing adults. They attributed this to the early manifestation of bilateral Meniere's disease. The findings in the present study also showed a similar trend, though for AP. The findings of the present studyindicate that the latency difference of AP between the polarities could be useful in the early detection of onset of bilateral Meniere's in the individuals who are already suffering from the unilateral condition.

ABR Wave I Latency Difference between Condensation and Rarefaction.

The results of the present study indicated towards a greater difference in Meniere's disease in terms of the latency difference between condensation and rarefaction click-evoked wave I of ABR. Tietze and Pantev (1986) reported a mean difference of 0.06 ms in ABR wave I latency between condensation and rarefaction polarities in normal hearing individuals. For high-level clicks, the wave I (Coats & Martin, 1977) or eighth nerve action potential (Peake & Kiang, 1962) has been reported to be approximately 0.2 ms earlier for rarefaction than for condensation stimuli. Contrary to this, Beattie (1988) found no significant differences between the latencies for human wave I elicited by rarefaction and condensation clicks. The findings of the present study are in agreement with those of Coats and Martin (1977) and in disagreement with Beattie (1988). However, these values were very small and the test-retest reliability of up to 0.3 ms has been reported (Hall, 2007) for different waves, including wave I of ABR. Thus such small deviations in normals could be ignored. The absence of or very small phase-related latency difference is expected for normal-hearing subjects because the normal click-evoked ABR is dominated by high-frequency neural responses (Don & Eggermont, 1978), which are not significantly affected by stimulus phase. However in the ears with Meniere's disease, the changes due to endolymphatic hydrops has been documented to result in prolongation of AP for condensation polarity, thereby by increasing the latency difference of AP between the two polarities (Saas et al., 1998; Chen et al., 2004). Since the wave I of ABR corresponds to the AP (Moller et al., 1983; Hall & Antonelli, 2001), a similar prolongation of latency for condensation polarity is justified. However, the present study is one of the first that has used ABR for the diagnosis of Meniere's disease and there are no other studies that reported about the usefulness of ABR in the diagnosis of Meniere's disease. The present study also revealed another intriguing finding. The unaffected ears of individuals with Meniere's disease revealed significantly larger difference between the latency of wave I of ABR between the polarities compared to the ears of healthy individuals and the ears with sensorineural hearing loss. This may be an early sign of progression of the disease to the unaffected ear in the individuals with unilateral Meniere's disease. There are a number of studies that have reported a tendency of Meniere's disease to progress to the other ear within 2-7 years of its onset in the first in more than 50% of the cases (Morrison, 1981; Salvinelli et al., 1998; Jackson & Silverstein, 2002; Saeed & Penny, 2004). But there are no studies that have used ABR wave I latency difference between the polarities and demonstrated such a finding. However, the reports using OAEs have shown that the unaffected ears of individuals with Meniere's disease had significantly lower amplitudes and the authors discussed this as an early sign of progression to bilateral condition (Huffelen et al., 1998). On similar lines the findings of present study may be considered an early sign for progression to bilateral Meniere's disease. So the present study shows that the technique of latency difference between condensation and rarefaction not only aids the diagnosis of Meniere's disease in the affected ears but also is capable of predicting its spread in the unaffected ear.

SP/ AP Ratio

The mean SP/AP amplitude ratios in the affected ears of individuals with Meniere's disease was higher when compared to ears of healthy individuals, ears of individuals with sensorineural hearing loss, and also the unaffected ears of individuals with Meniere's disease in the present study. The mean SP/AP ratios in these were 0.26, 0.24, 0.26, and 0.43 respectively. Similar patterns of findings have been reported in literature about SP/AP ratio (Mori et al., 1987; Chen et al., 2004). Mori et al. (1987) conducted a study in which they found the mean SP/AP ratio of 0.22, 0.20, and 0.63 respectively in ears of healthy individuals, ears of individuals with sensorineural hearing loss, and ears with Meniere's disease. Chen et al. (2004) obtained a mean SP/AP ratio of 0.46 in the individuals with Meniere's disease and 0.22 in the ears of healthy individuals. The differences between the ratio values observed in the present study and that by Mori et al. (1987) may be because of the differences in the method and also the presence of a higher variability within the group with Meniere's disease in terms of higher standard deviation for the Meniere's group in their study. The present study used Tiptrode placed at the entrance of the ear canal as the non-inverting electrode as against silver ball electrode placed near the tympanic membrane at a distance of 3 mm in the study by Mori et al. (1987). Also, the standard deviation obtained was 0.19 in the present study as opposed to 0.44 in Mori et al. (1987). Furthermore, the volatile nature of the Meniere's disease itself, which causes high variability in the test results usually, may have contributed to the differences between their study and the present study. However, Chen et al. (2004) used the same electrode placement (extratympanic) as the present study and obtained a mean SP/AP ratio of 0.46 in Meniere's disease and 0.22 in normal hearing individuals. These findings are similar to the ones observed in the present study and thus it probably extends further support to reasons of the differences observed between the present study and the study by Mori et al. (1987).

Relationship between AP Latency Difference and SP/AP Ratio in the Affected Ears of the Individuals with Meniere's Disease.

The present study revealed the existence of *no significant* correlation between AP latency difference and SP/AP ratio. Similar findings have also been reported by Ohashi et al. (2009). This disagreement between the techniques may be attributed to the differences in the physiology of the AP and SP generation. The AP waveform is characterized by a series of brief, predominantly negative peaks representing the distribution of underlying neural firings. The response to moderately intense stimulation tends to be dominated by neural contributions from the basal or high frequency end of the cochlea (Kiang, 1965), at least in normal ears and pathologic ears no worse than moderate hearing loss. The AP magnitude can also be viewed as a reflection of inner hair cell output. The SP is a complex response made up of several components. It reflects the displacement time pattern of cochlear partition. The SP is stimulus related and generated by the hair cells of the organ of Corti (Dallos, 1973). The SP manifests itself as a shift in the CM baseline, the direction of which is indicated and dictated by an interactive effect between stimulus parameters and the location of the recording electrode (Dallos, 1973). In general, the ECochG waveform recorded from patients with suspected endolymphatic hydrops is often characterized by an enhanced summating potential (Ferraro, Arenberg & Hassanein, 1985).The rationale usually given for this finding is that an increase in endolymph volume alters the hydromechanical characteristics of the inner ear because of the resultant increase in intra-labyrinthine pressure. When this occurs, the normal vibratory asymmetry of the basilar membrane is augmented. Since the SP supposedly reflects this vibratory asymmetry, it too will be enhanced during a hydropic state. However, AP has a distinctly different physiology and hence the presentation in hydropic pathology may accordingly be different. This probably may explain the disagreement between the two techniques for the diagnosis of MD and calls for further studies to clarify the reasons.

In the present study the agreement between the two measures was found only for 9 ears. This implied that a Meniere's disease versus non-Meniere's disease diagnosis was appropriately made in only 56.25 % of the individuals when using a criterion of positive results on both techniques. However, when the criterion was changed to positive result on either of the two, 85 % of the cases were diagnosed as Meniere's disease. So both these techniques should be used in the protocol for the diagnosis of Meniere's disease as the two are likely to give complementary information, thereby supplementing in the diagnosis. It might be interesting to note if there would be a correlation between the agreement of the two tests and the stage of the Meniere's disease diagnosed as per AAO-HNS (1995). However, due to smaller sample size, a correlation of this kind could not be taken up. This may be considered in future studies using the two techniques and if found to correlate, it could well become an objective way of staging Meniere's disease.

Relationship between ABR Wave I Latency Difference and SP/AP Ratio in the Affected Ears of the Individuals with Meniere's Disease

The results of the present study revealed a lack of relationship between ABR wave I latency difference and SP/AP ratio in the affected ears of the individuals with Meniere's disease. Lack of agreement may be due to the differences in the physiology of generation of wave I of ABR and SP. Wave I of ABR is generated at the distal part of the auditory nerve (Wada & Starr, 1983) whereas SP is generated from the hair cells in the cochlea (Dallos, 1973). The two are likely to represent different aspects of physiology owing to these different structures (hair cells versus nerve fibers) involvement. Further, there are no such studies in the literature that compare latency difference between the two polarities for wave I of ABR and SP/AP in individuals with Meniere's disease. Present study is one of the first of this kind and the findings indicate towards a lack of correlation between this technique and the more established SP/AP amplitude ratio in the diagnosis of Meniere's disease. However, wave I of ABR has been reported to correspond to AP of ECochG (Moller & Janetta, 1983; Hall& Antonelli, 2001) and a similar lack of correlation of AP with SP/AP ratio (Ohashi et al., 2009) could explain the results of the lack of correlation so found.

Relationship between AP Latency Difference and ABR Wave I Latency Difference in the Affected Ears of Individuals with Meniere's Disease

The present study indicated a *strong correlation* between AP latency difference and ABR wave I latency difference in the affected ears of individuals with Meniere's disease. Tonndorf (1976), through his cochlear model, postulated that the basilar membrane undergoes a downward displacement when loaded with an excessive amount of endolymph. Distortion at the level of the basilar membrane will generally be reflected at the auditory nerve fibers in terms of the neural responses. Since wave I of ABR reflects the activity in the distal portion of the auditory nerve (Wada & Starr, 1983) and also corresponds to AP (Moller & Janetta, 1983, Hall & Antonelli, 2001), the finding of a strong correlation is expected and justified.There are no such studies in the literature that compare the latency difference of wave I of ABR and AP latency difference between the two polarities. So, the findings of the present study support the utility of latency difference between the polarities in the diagnosis of Meniere's disease, atleast to the same degree as that of AP latency difference between the polarities.

Conclusions

The latency difference of AP between condensation and rarefaction polarities can be an useful tool for the diagnosis of Meniere's disease. A similar difference for ABR wave I can also be used effectively for its diagnosis. The techniques of AP latency difference and ABR latency difference are better suited to the diagnosis of Meniere's disease than SP/AP amplitude ratio. However a larger data point needs to be used for the further validation of the results. Additionally, a combination of SP/AP ratio and AP latency shift would increase the chance of a positive diagnosis of Meniere's disease.

References

American Academy of Otolaryngology-Head and Neck Surgery Committee on Hearing and Equilibrium Committee on hearing and equilibrium guidelines for the diagnosis and evaluation of therapy in Meniere's disease (1995). *Otolaryngology- Head and Neck Surgery, 113*, 181-185.

American National Standards Institute (1999). *American National Standards for Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms*. ANSI S3.1 (1999). New York: American National Standards Institute.

Beattie, R. C. (1988). Interaction of click polarity, stimulus level, and repetition rate on the auditory brainstem response.*Scandinavian Audiology, 17,* 99-109.

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

Chen, Y. C., Kang, B. H., Yeh, W. Y., & Wang, H. W. (2004).Electrocochleography for normal adults and patients of Meniere's disease.*Journal of Medical Science, 24*(6), 313-318.

Coats, A. G., & Martin, J. L. (1977). Human auditory nerve action potential and brainstem evoked responses. Effects of audiogram shape and lesion location.*Archives of Otolaryngology, 103,* 605-622.

Cullen, J. K., Ellis. M. S., & Berlin, C. I. (1972). Human nerve action potential recordings from the tympanic membrane without anesthesia. *Acta Otolaryngologica, 74,* 15-22.

Dallos, P. (1973). *The auditory periphery: biophysics and physiology.* New York: Academic Press.

Don, M. & Eggermont, J. J. (1978). Analysis of click evoked brainstem potentials in man using high pass noise masking. *Journal of the Acoustical Society of America, 63,* 1084-1092.

Eggermont, J. J., & Odenthal, D. W. (1974).Frequency selective masking in electrocochleography. *Revised Laryngology Oto Rhinology, 95,* 489-495.

Eggermont, J. J., Odenthal, D. W., Schmidt, P. M., & Spoor, A. (1974). Electrocochleography: basic principles and clinical application. *ActaOtolaryngologica, 316*.

Ferraro, J. A., Arenberg, I. K., & Hassanein, R. S. (1985). Electrocochleography and symptoms of inner ear dysfunction. *Archives of Otolaryngology, 111,* 71-74.

Hall, J. W., & Antonelli, P. J. (2001). Assessment of peripheral and central auditory function. In B. J. Bailey, R. K. Jacler, H. C. Pillsbury & P. R. Lambert (3rd edition), *Head and Neck Surgery Otolaryngology*. Philadelphia, USA: Lippincot, Williams & Wilkins.

Hall, J. W. (2007). *New handbook of auditory evoked responses.* Boston, MA: Pearson Education Inc.

Huffelen, W. M., Mateijsen, N. J. M., & Wit, H. P. (1998). Classification of patients with Meniere's disease using otoacoustic emission.*Audiology and Otoneurology, 3,* 419-430.

Jackson, L. E. & Silverstein, H. (2002). Chemical perfusion of the inner ear. *Otolaryngologic Clinics of North America, 35,* 639-653.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., & Clark, L. F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve.* Cambridge, MA: The MIT Press.

Kimura, R. S. (1982). Animal models of endolymphatic hydrops. *American Journal of Otolaryngology, 3,* 447-451.

Margolis, R. H., Rieks, D., Fournier, E. M., & Levine, E. M. (1995). Transtympanic electrocohleography for diagnosis of Meniere's disease. *Archives of Otolaryngology-Head and Neck Surgery, 121,* 44-55.

Moffat, D. A., Gibson, W. P. R., Ramsden, R.T., Morrison, A. W., & Booth, J. B. (1977). Transtympanic electrocochleography during glycerol dehydration. *Acta Otolaryngologica, 85,* 158-166.

Moller, A. & Janetta, P. (1983). Monitoring auditory functions during cranial nerve microvascular decompression operations by direct monitoring from the eight nerve. *Journal of Neurosurgery, 59,* 493-499.

Mori, N., Asai, A., Suizu, Y., & Ohta, K. (1985). Comparison between electrocochleography and glycerol test in the diagnosis of Meniere's disease. *Scandinavian Audiology, 14,* 209-213.

Mori, N., Asai. H., & Matsunaga, T. (1987).Diagnostic value of extratympanic electrocochleography in Meniere's disease.*International Journal of Audiology, 26*, 103-110.

Morrison, A. W. (1981). Meniere's disease. *Journal of the Royal Society of Medicine, 74,* 183Â18.

Ohashi, T., Nishino, H., Arai, Y., Hyodo, M., &Takatsu, M. (2009).Clinical significance of the summating potential- action potential latency difference for condensation and rarefaction clicks in Meniere's disease. *Annals of Otorhinolaryngology, 118*(4), 307-312.

Orchik, D. J., Ge, N. N., & Shea, J. J. (1998). Action potential latency shift by rarefaction and condensation clicks in Meniere's disease. *Journal* *of the American Academy of Audiology, 9*(2), 121-126.

Peake, W. T., & Kiang, N. Y. (1962). Cochlear responses to condensation and rarefaction clicks. *Journal of Biophysics, 2,* 23.

Prijs, V. F. (1986). Single-unit response at the round window of the guinea pig. *Hearing research, 21,* 127-133.

Roland, P. S., Yellin, M. W., Meyerhoff, W. L., & Frank, T. (1995). Simultaneous comparison between transtympanic and extratympanic electrocochleography. *American Journal of Otolaryngology, 16*(4), 444-50.

Saeed, S., & Penny, S. (2004). Diagnosis and management of Meniere's disease. *ENT News,13,* 32-34.

Salvinelli, F., Trivelli, M., Greco, F., Silvestrini, M., Fernandez, E., & Pallini, R. (1999). Meniere's disease: is it a bilateral disease? *European Review for Medical and Pharmacological Sciences, 3,* 129-133.

Santarelli, R., & Arslan, E. (2002). Electrocochleography in auditory neuropathy. *The Hearing Research, 170*(1-2), 32-47.

Sass, K., Densert, B., Magnussen, M., & Whitaker, S. (1998). Sensitivity and specificity of transtympanicelectrocochleography in Meniere's disease. *ActaOtolaringologica, 118*(2), 150-156.

Tietze, G., & Pantev, C. (1986). Comparison between auditory brainstem responses evoked by condensation and rarefaction step functions and click. *Audiology, 25,* 44-53.

Tonndorf, J. (1976). Endolymphatic Hydrops: Mechanical causes of Hearing loss. *Archives of Oto-Rhino-Laryngology, 212,* 293- 299.

van Deleen, G. W., & Smoorenburg, G.F. (1986). Electrocochleography for different electrode positions in guinea pig. *ActaOtolaringologica, 101,* 207-216.

Viera, A. S., & Garrett, J. M. (2005). Understanding inter observer agreement: the Kappa statistics. *Research Series, 37*(5), 360-366.

Wada, S. I. & Starr, A. (1983).Generation of auditory brainstem responses (ABRs). I. Effects of injection of a local anesthetic (procaine HCl) into the trapezoid body of guinea pigs and cat. *Electroencephalography and Clinical Neurophysiology, 56*, 326-339.

Stimulus Rate and Subcortical Auditory Processing of Speech: Comparison between Younger and Older Adults

[1]Garvita & [2]Sinha S. K.

Abstract

Perception of acoustic signal depends on accurate encoding of temporal events of auditory signals. The auditory brainstem reflects processing of temporal events those are diagnostically significant in the assessment of hearing loss and neurological function (Hall, 1992). Degenerative changes occur with advancing age in the central auditory pathway, including both sub cortical and cortical structures. 17 young aged individuals (30 ears- 18 to 30 years) and 15 older adults (30 ears- 40 to 55 years) with normal hearing sensitivity participated in the study. The brainstem responses to click stimulus was recorded presented at 80 dB SPL across three repetition rates) and was analyzed for wave V latency. The brainstem response to speech evoked ABR was also recorded by using syllable /da/ presented at 80 dB SPL across the three repetition rates (6.9, 10.9 & 15.4). The speech ABR waveforms were analyzed for both the onset (latency of wave V and A) and the sustained (latencies of wave - D, E and F) responses. FFT was done to find the raw amplitude of F0, F1 and higher harmonics (F2) frequency components. The increase in latency of speech evoked ABR (transient and sustained response) and click evoked ABR for the older adults suggest that the brainstem timing might be affected for the older adults. A reduction in amplitude for the coding of F0, F1 and F2 for the older adult group was seen. Reduction in coding of F0, F1 and F2 might be leading to the speech perception problems in older individuals.

Keywords: *Speech ABR, older adults,, repetition rate*

Introduction

Older aged individuals have been shown to have greater difficulty with speech understanding than younger listeners (CHABA, 1988). However, there are studies which demonstrate that in adverse listening conditions, older individuals with essentially normal peripheral hearing sensitivity, have difficulty in understanding speech (Ewertsen & Birk-Nielsen, 1971; Plomp & Mimpen, 1979; Nabelek & Robinson, 1982; Era, Jokela, Qvarnberg & Heikkinen, 1986; Gelfand, Piper & Silman,1986; Dubno, Horwitz & Ahlstrom, 2002; Kim, Frisina, Mapes, Hickman & Frisina, 2006; Wingfield, McCoy, Peelle, Tun & Cox, 2006). This may lead one to conclude that age-related changes occur beyond the peripheral auditory system, i.e., the central auditory nervous system might play a role in this difficulty (Gordon-Salant, 1987; Humes, 1996; Frisina & Frisina, 1997; Mazelova, Popelar & Syka, 2003). These studies have a group of subjects in the middle age range i.e. in the age range of 40-60 years.

Studies suggest that certain auditory abilities begin to decline in older adult population. For example, Barr and Giambra (1990) reported that middle-aged subjects perform more poorly than younger listeners (but better than older individuals) on tasks such as perception of dichotically presented speech. Bergman (1971) reported a significant decline in perception of interrupted speech in middle aged individuals. As the difficulty in speech understanding in elderly individuals arises from the central auditory nervous system, the decline of speech understanding in older adults also may arise from the central auditory nervous system itself. One form of central auditory processing that has been attributed to part of this difficulty in older elderly individuals is the temporal processing (Fitzgibbon & Gordon-Salant, 1996).

Temporal processing is one of the functions necessary for the discrimination of subtle cues such as voicing and discrimination of similar words. Deficits of temporal processing have been found in a group of 40to55 year-old individuals (Grose, Hall & Buss, 2006), reduced gap detection ability in middle-aged women (Helfer & Vargo, 2006), subtle differences in auditory perception between younger and older adults in auditory event-related potential (Alain, McDonald, Ostroff & Schneider, 2004; Geal-Dor, Goldstein, Kamenir & Babkoff, 2006) processing of interaural phase differences both in behavioral and physiological tasks (Ross, Fujioka, Tremblay & Picton, 2007) demonstrating that age-dependent subtle auditory changes may begin in older adults. One way to assess the temporal processing electrophysiologically is to study the stimulus complexity by examining the effects of stimulus rate on speech evoked auditory brainstem responses (Krizman, Skoe & Kraus, 2010; Basu, Krishnan & Weber-Fox, 2010). Click-evoked ABR is a gross measure of time-locked neural activity in response to stimulus onset. In contrast, the frequency-following response (FFR) is a steady state AEP that is sensitive to sustained features within a stimulus and is dependent on the integrity of phase-locked neural activity in the auditory brainstem (Worden & Marsh, 1968).

By increasing the repetition rate of the stimuli, the au-

[1]Email: garvita28mehta@gmail.com,
[2]Lecturer in Audiology, Email: sujitks5@gmail.com

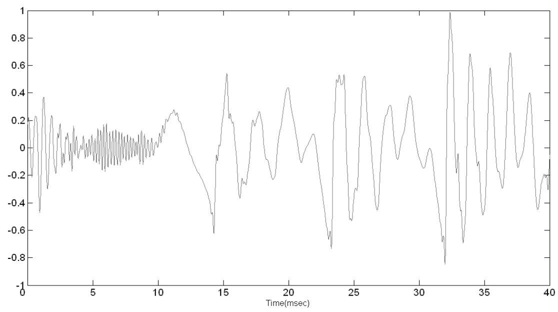


Figure 1: Time domain waveform the Speech stimulus /da/ used in the present study. The top one represents the temporal details of the waveform whereas the bottom one depicts the spectral details.

ditory temporal processing can be checked. By utilizing speech stimulus for assessing the temporal processing will give additional information about temporal coding of speech at the brainstem. Present study was taken up with an aim of investigating the interactions between auditory temporal processing and stimulus complexity by examining the effects of stimulus rate on speech evoked and click evoked ABR in normal hearing younger adults and older adults and to check whether the stimulus rate affects the encoding of the onset of the response or the sustained portion of the response in older adults.

Method

Participants

Younger adults- 17 participants (30 ears) in the age range of 18 to 30 years (Mean age= 21.8 years) and older adults - 15 participants (30 ears) with normal hearing sensitivity in the age range of 40 to 55 years (Mean age= 47.3 years) participated in the study.

Instrumentation and Test Environment

Pure Tone Audiometry was done to confirm bilateral normal hearing sensitivity. Immittance audiometry was done to rule out middle ear abnormalities. Biologic Navigator Pro EP was used to record both click evoked and speech evoked ABR. All the audiological evaluation and recording were carried out in a sound treated room. The ambient noise was within the permissible limits as recommended by ANSI (S3.1; 1991).

Test Stimulus for Speech ABR

Figure 1 shows both the time and spectral domain of the stimulus used in the present study. The stimulus is available in evoked potential system with the BioMARK protocol. The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980). This stimulus simultaneously contains

Table 1: Recording protocol of the click and speech evoked ABR

Parameters	Click evoked ABR	Speech evoked ABR
Stimulus, duration	Click, 100 μs	CV syllable /da/, 40 ms
Level	80 dB SPL	80 dB SPL
Filter band	100 to 3000 Hz	100 to 3000 Hz
Rate	9.1/s, 19.1/s & 40.1/s	6.9/s, 10.9/s & 15.4/s
No of sweeps	2000	2000
Transducer	BioLogic Insert ear phone	BioLogic Insert ear phone
Polarity	Alternating	Alternating
Time window	12 ms	64 msec which included a prestimulus time of 10 ms (default setting in Biologic system)
Electrode montage	Non-inverting electrode:ForeheadInverting electrode: Test ear MastoidGround electrode: Non test ear mastoid.	Non-inverting electrode:ForeheadInverting electrode: Test ear Mastoid.Ground electrode: Non test ear mastoid.

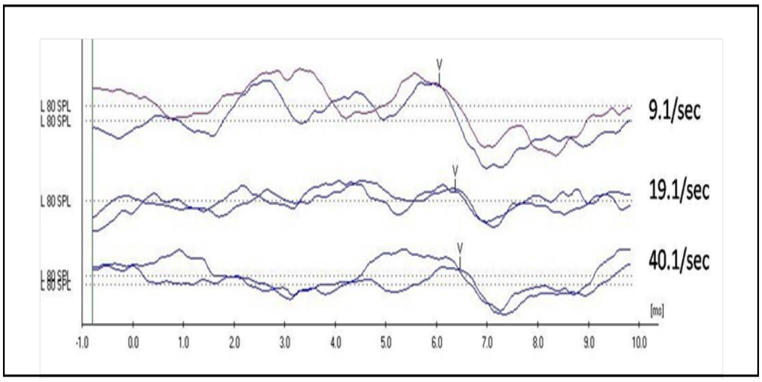


Figure 2: A sample waveform of click evoked ABR recorded at three different repetition rates in a young adult.

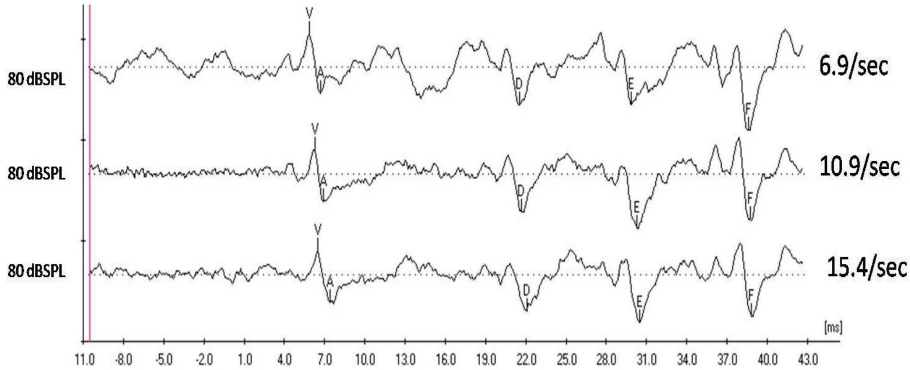


Figure 3: A sample waveform of Speech evoked ABR- transient and FFR waveform at three repetition rates obtained from one young group individual.

broad spectral and fast temporal information characteristics of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. The fundamental frequency (F0) linearly rises from 103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 msec. The first formant (F1) rises from 220 to 720 Hz, while the second formant (F2) decreases from 1700 to 1240 Hz over the duration of the stimulus. The The third formant (F3) falls slightly from 2580 to 2500 Hz, while the fourth (F4) and fifth formants (F5) remain constant at 3600 and 4500 Hz, respectively.

Procedure

Click ABR and Speech ABR was recorded using the protocol shown in Table 1.

Results

Effect of Repetition Rate and Age (young and older adult group) on the Latency of Click Evoked ABR

The latency of wave V was analyzed for the click evoked ABR across the three different repetition rates (9.1, 19.1 & 40.1/sec). Figure 2 shows an ABR waveform elicited by click at three repetition rates in normal hearing young adult.

It can be seen in the figure 2 that there is an increase in latency of the click evoked wave V as the repetition rate increased. To see the effects of repetition rate on latency of click evoked ABR wave V, Mixed ANOVA was done. Mixed ANOVA revealed a significant main effect for three repetition rates [F (2,116) = 126.11, p<0.05], but Mixed ANOVA failed to show any significant interaction between the repetition rates and the two groups [F (2,116) = 1.94, p>0.05]. Mixed ANOVA also showed a significant difference for the two groups [F (1, 58) = 40.77, p<0.05]. Bonferroni pairwise comparison test was done to see the groupwise differences for the three repetition rates. Pairwise comparison test showed a significant difference between the repetition rate 9.1- 19.1,

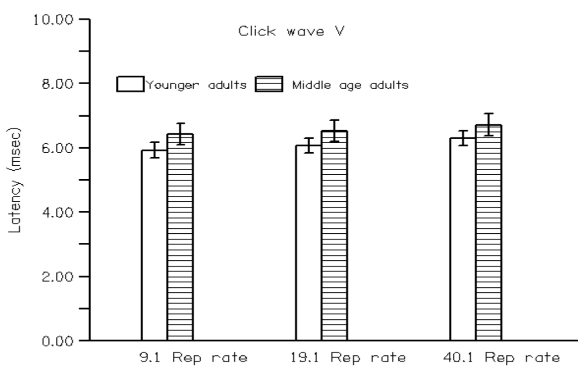


Figure 4: Mean latency of wave V peak latency of click ABR for three repetition rates across the two groups.

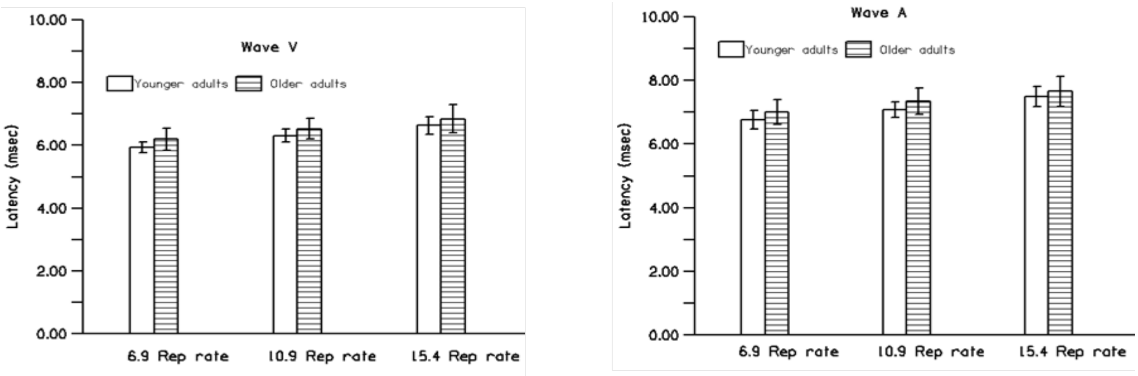


Figure 5: Latency of speech evoked transient V and A peak latency for three repetition rates across the two groups.

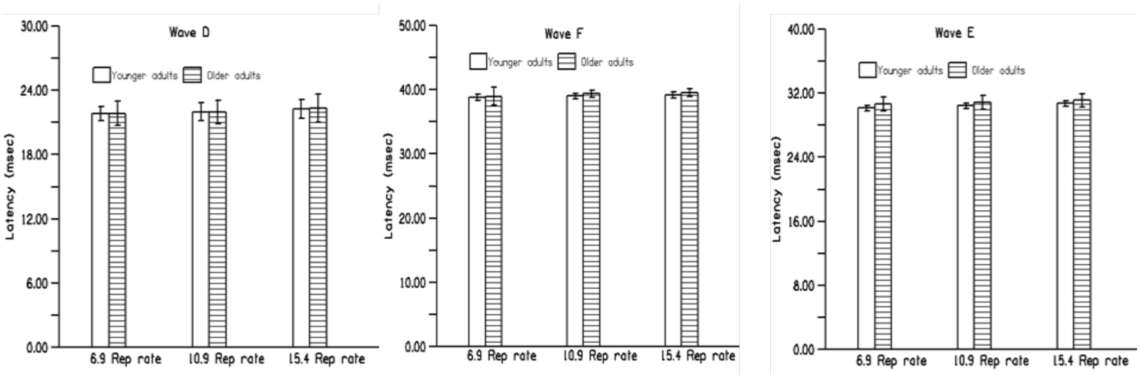


Figure 6: Latency of speech evoked D, E and F peak latency for three repetition rates across the two groups.

9.1- 40.1 and 19.1- 40.1/sec.

As the Mixed ANOVA showed significant interaction across the repetition rates, taking data from both the groups, Repeated measure ANOVA test was done which showed that there was significant difference across the three repetition rates for the young group [F (2, 58) = 66.71, p<0.05] and for the older adult group [F (2, 58) = 60.18, p<0.05]. Thus, Bonferroni post hoc test was done which revealed a significant difference between the repetition rate 9.1- 19.1, 9.1- 40.1 and 19.1- 40.1/sec for young and older adult group.

Effect of Repetition Rate and Age (young and older adult group) on the Latency of Speech Evoked ABR

The latency of wave V, A, D, E, F was analyzed for the speech evoked ABR for three different repetition rates (6.9, 10.9, and 15.4). Figure 3 shows syllable /da/ evoked ABR and FFR waveform at three repetition rates obtained from one of the young group individual.

As it can be seen in the Figure 3, that there is an increase in latency of all the peaks of speech evoked transient response (wave V and wave A) and FFR (wave D, E and F) with the increase in repetition rate. The mean and standard deviations for the latency of different peaks of speech evoked auditory brainstem responses were calculated.

Mixed ANOVA revealed a significant main effect for repetition rates for the latency of wave V, A D, E and F of speech evoked ABR [F (2, 116) = 155.34, p< 0.05], but Mixed ANOVA failed to show any significant interaction between repetition rates and groups [F (2, 116) = 0.17 p> 0.05]. Mixed ANOVA revealed a significant difference between the two groups [F (1, 58) = 4.78, p< 0.05]. Bonferroni pairwise comparison was done to see the group wise differences for the three repetition rate which revealed significant difference between 6.9- 10.9, 6.9- 15.4 and 10.9- 15.4/ sec repetition rate.

Results of Onset responses (Wave V and Wave A): As the Mixed ANOVA showed significant interaction across the repetition rates, taking data from both the groups, Repeated Measure ANOVA (3 repetition rates) was done within the group, results showed that for wave V latency, there was a significant difference across the three repetition rates for younger group [F(2, 58)= 292.93, p< 0.05] and the older adult group [F(2, 58)= 169.50, p< 0.05] and for wave A latency also, there was a significant difference across the three repetition rates for younger group [F(2, 58)= 115.05, p< 0.05] and the older adult group [F(2, 58)= 107.85, p< 0.05]. Thus, Bonferroni post hoc test was done to see, at which two

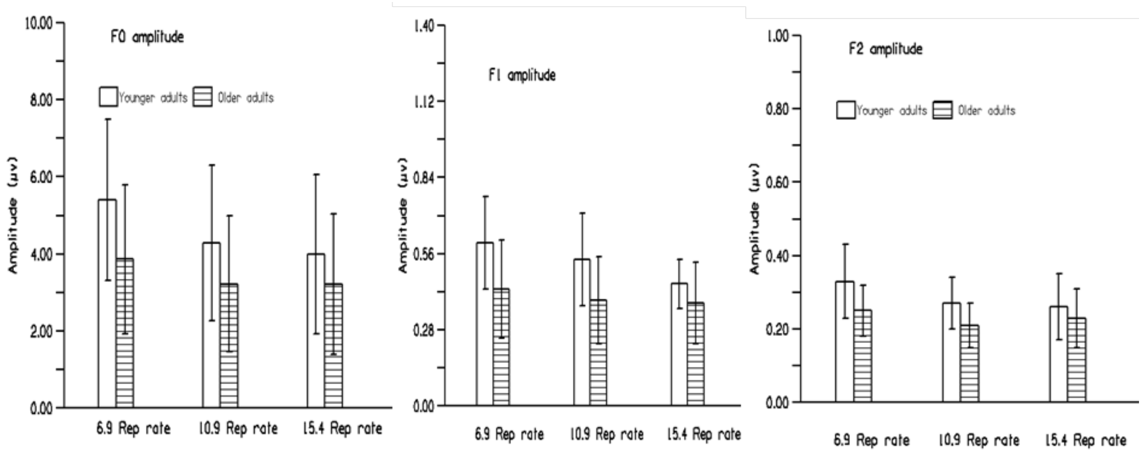


Figure 7: Amplitude of F0, F1 and F2 at 6.9, 10.9 and 15.4 repetition rates across the two groups.

repetition rates, latency differed significantly for wave V and wave A, results revealed a significant difference 6.9- 10.9, 6.9- 15.4 and 10.9- 15.4/ sec repetition rate for both the groups.

*Results of Sustained Responses (Wave D, E and F):*As the Mixed ANOVA showed significant interaction across the repetition rates, taking data from both the groups, Repeated Measure ANOVA (3 repetition rates) was done for wave D, results showed that there is a significant difference across the three repetition rates for younger group [F (2, 58) = 18.39, p< 0.05] and the older adult group [F (2, 58) = 9.22, p< 0.05]. Thus, Bonferroni post hoc test was done for wave D, results obtained showed a significant difference between 10.9- 15.4 and 6.9- 15.4/sec repetition rate for both the groups.

Similarly Repeated Measure ANOVA (3 repetition rates) was done for wave E. Test results showed that there is a significant difference across the three repetition rates for younger group [F(2, 58)= 40.14, p< 0.05] and the older adult group [F(2, 58)= 10.62, p< 0.05]. Thus, Bonferroni post hoc test was done for wave E and the results obtained showed a significant difference between all the repetition rates except between 6.9- 10.9 for the older age group.

Similarly Repeated Measure ANOVA (3 repetition rates) was done for wave F. Test results showed that there is a significant difference across the three repetition rates for the younger group [F(2,58)=44.33, P<0.05] and the older group [F(2,58)=3.39, P<0.05]. Boneferroni post hoc test showed a significant difference between 6.9-15.4 and 10.9-15.4 for the young group only.

Since, the wave V latency increased with increase in repetition rate, the delay in sustained response (wave D, E & F) might be due to delay in wave V. To understand whether the delay in the sustained response was because of increase in latency of wave V or repetition rate affected the sustained reponse latency, the wave D, E & F latency were covaried for 10.9 & 15.4 repetition rate with respect to 6.9/sec.

For wave D, a Repeated Measure ANOVA showed a significant difference across the repetition rates for the young group [F(2, 58)= 11.87, p< 0.05] but failed to show a significant difference for the older adult group [F(2, 58)= 1.71, p> 0.05] and thus, Bonferroni post hoc test was done for the young age group and a significant difference was found between 6.9/sec and 10.9/sec (p<0.05), between 6.9/sec & 15.4/sec (p<0.05) but not between 10.9/sec & 15.4/sec (p>0.05).

For wave E, a Repeated Measure ANOVA failed to show any significant difference across the repetition rates for the young group [F (2, 58)= 1.97, p> 0.05] and for the older adult group [F(2, 58)= 2.41, p> 0.05].

For wave F, a Repeated Measure ANOVA showed a significant difference across the repetition rates for the young group [F(2, 58)= 26.62, p< 0.05] but failed to show any significant difference for the older adult group [F(2, 58)= 0.43, p> 0.05]. Thus, Bonferroni post hoc test was done for the young age group and a significant difference was found between 6.9/sec and 10.9/sec (p<0.05), between 6.9/sec & 15.4/sec (p<0.05) and between 10.9/sec & 15.4/sec (<0.05).

Effect of Repetition Rate and Age (Young and Older Adult Group) on the Amplitude of Pitch and Harmonics of Speech Evoked ABR

The amplitude of F0 (103 to 125 Hz), F1 (220 to 720 Hz) and higher harmonics (F2- 1700 to 1240 Hz) was analyzed for the speech evoked ABR for three different repetition rates (6.9, 10.9, 15.4/s). The mean and the SD of the amplitude of F0, F1 and F2 were calculated for the speech evoked FFR recorded at three repetition rates. To see the significant difference across three repetition rates on F0, F1 and F2 mean amplitude, Mixed ANOVA (3 repetition rates and 2 groups) was done. Mixed ANOVA revealed a significant main effect across the repetition rates [F (2, 116) = 7.58, p< 0.05], but Mixed ANOVA failed to show any significant interaction between the repetition rates and the two groups [F (2, 116) = 1.01, p> 0.05]. Mixed ANOVA revealed a significant difference across two groups [F (1, 58) = 13.66, p< 0.05]. Thus, Bonferroni pairwise comparison was done to see the group wise differences for the three repetition rates and the results revealed as significant difference between 6.9- 10.9 and between 6.9- 15.4/sec repetition rate.

Results of Amplitude of F0- Pitch measure: As the Mixed ANOVA showed significant interaction across the repetition rates, Repeated Measure ANOVA (3 repetition rates) was done within the groups, results showed that there is a significant difference across the three repetition rate for younger group [F(2, 58)= 4.93, p< 0.05], but failed to show a significant difference for the older adult group [F(2, 58)= 1.40, p> 0.05]. Thus, the young group, Bonferroni post hoc test was done which revealed a significant difference between 6.9- 15.4/ sec repetition rate only.

Amplitude of F1: As the Mixed ANOVA showed significant interaction across the repetition rates, Repeated Measure ANOVA (3 repetition rates) was done within the groups, results showed that there is a significant difference across the three repetition rates for younger group [F(2, 58)= 20.41, p< 0.05] but not in the older adult group [F(2, 58)= 1.88, p> 0.05]. Thus, the young group, Bonferroni post hoc test was done which revealed a significant difference between 6.9- 10.9, 6.9- 15.4 and 10.9- 15.4/sec repetition rate.

Amplitude of F2: As the Mixed ANOVA showed significant interaction across the repetition rates, Repeated Measure ANOVA (3 repetition rates) was done within the groups, results showed that there is a significant difference across the three repetition rates for younger group [F(2, 58)= 14.15, p< 0.05] and in the older adult group [F(2, 58)= 4.87, p<0.05]. Thus, Bonferroni post hoc test was done which revealed a significant difference between 6.9- 10.9, 6.9- 15.4/sec for young group and for 6.9- 10.9/sec repetition rate for the older adult age group.

Latency of Click ABR across the Young and Older Adults

Figure 4 shows the mean latency of wave V peak latency of click ABR for three repetition rates across the two groups.It can be seen from the Figure 4 that at all the three repetition rates, the mean latency of wave V of click ABR for older adult group was more prolonged than the young age group. Multiple analysis of variance (MANOVA) was done to understand the significant difference in latency of wave V of click ABR for the two groups across the three repetition rates. MANOVA results showed significant difference in wave V latency of click ABR across the two groups for repetition rate 9.1 [F (1, 58) = 44.013, p< 0.05], 19.1 [F (1, 58) = 36.718, p< 0.05], 40.1 [F (1, 58) = 30.169, p< 0.05].

Latency of Speech Evoked ABR Across the Young and Older Adult Groups

Latency of onset response: The following Figure 5 shows the mean latency of speech evoked transient V and A peak latency for three repetition rates across the two groups. It can be seen from the figure 5, that at all the three repetition rates, the latency of wave V and A for older adult group was more prolonged than the young age group. For wave V latency, MANOVA results showed significant difference across the groups for repetition rate 6.9 [F (1, 58) = 13.70, p< 0.05], 10.9 [F (1, 58) = 9.96, p< 0.05], 15.4 [F(1, 58) = 4.83, p< 0.05]. For wave A latency, MANOVA results showed significant difference across the groups for repetition rate 6.9 [F(1, 58)= 7.40, p< 0.05], 10.9 [F(1, 58)= 8.92, p< 0.05] , but no significant difference for repetition rate 15.4 [F(1, 58)= 2.14, p> 0.05].

Latency of sustained response of speech evoked ABR across the groups: The following figure 6 shows the mean latency of D, E and F peak for three repetition rates across the two groups. It can be seen from the Figure 6, that at all the three repetition rates, the mean latency of wave D for older adult group is almost similar to the young age group and the mean latency of wave E and F for older adult group was more prolonged than the young age group.

Multiple analysis of variance was done to see the significant difference in latency of wave D, E and F across the young and the older adult groups. MANOVA results showed no significant difference in wave D latency across the groups for repetition rate 6.9 [F (1, 58)= 0.005, p> 0.05], 10.9 [F(1, 58)= 0.001, p> 0.05] and 15.4 [F(1, 58)= 0.097, p> 0.05]. MANOVA results showed significant difference in wave E latency across the groups for repetition rate 6.9 [F (1, 58) = 9.89, p< 0.05], 10.9 [F (1, 58) = 7.23, p<0.05] and 15.4 [F(1, 58)= 5.98, p< 0.05]. MANOVA results showed no significant difference in wave F latency across the groups for repetition rate 6.9 [F (1, 58) = 0.18, p> 0.05], but showed significant for 10.9 [F (1, 58) = 9.197, p<0.05] and 15.4 [F (1, 58) = 5.250, p< 0.05].

Latency of wave D, E and F were covaried for 10.9 & 15.4 repetition rates with respect to 6.9/sec repetition rate in order to understand the significant difference in wave D, E and F was due to wave V prolongation or there was an actual delay in wave D, E and F at higher repetition rate. Multiple analysis of variance was done with the covaried values of 10.9 and 15.4 repetition rates with respect to the latency of wave V of 6.9 repetition rate. MANOVA results showed no significant difference in wave D latency for the 10.9/sec [F (1, 58) = 0.02, p>

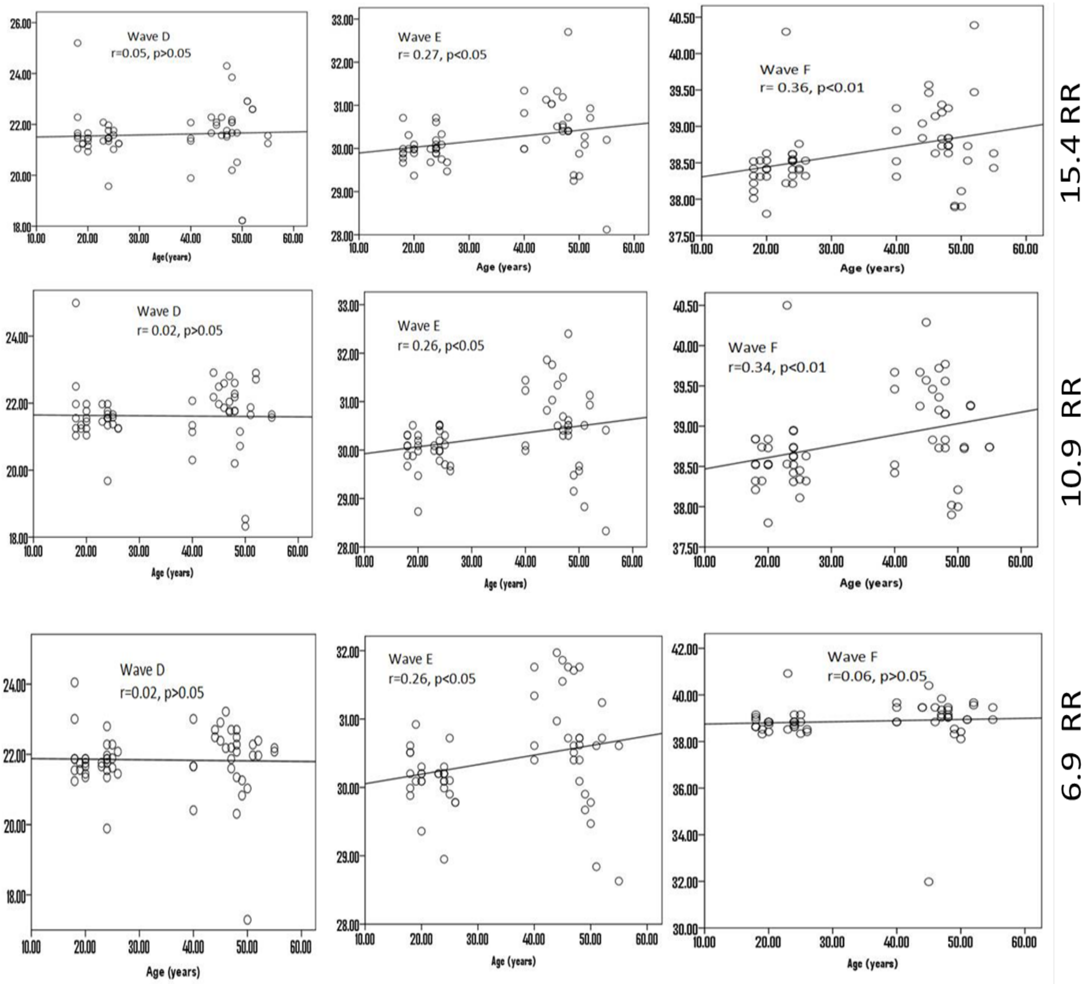


Figure 8: Correlation analysis between the age and the latency of the frequency following responses (sustained responses).

0.05] and for 15.4/sec [F (1, 58) = 0.27, p> 0.05] between the young and older adult groups for the covaried values. MANOVA results showed a significant difference in wave E latency for the 10.9/sec [F (1, 58) = 7.78, p< 0.05] and for 15.4/sec [F (1, 58) = 7.89, p< 0.05]. MANOVA results also showed significant difference in wave F latency for the 10.9/sec [F(1, 58)= 10.36, p< 0.05] and for 15.4/sec [F(1, 58)= 8.84, p<0.05].

Amplitude of the F0, F1 and Higher Harmonics (F2) across the Young and Older Adult Groups

The mean amplitude of F0, F1 and F2 comparing between young and older adult group is shown in the Figure 7.It can be seen from the Figure 7, that at all the three repetition rates, the F0, F and F2 amplitude of older adult group was lesser than young age group. Multiple analysis of variance was done to understand the significant difference for the amplitude of F0 between the young and the older adult group. MANOVA results showed significant difference in F0 amplitude across the groups for repetition rate 6.9 [F(1, 58)= 8.902, p<0.05], 10.9 [F(1, 58)= 4.799, p<0.05] , but showed no significant difference for repetition rate 15.4 [F(1, 58)= 2.402, p> 0.05]. MANOVA results showed significant difference in F1 amplitude across the groups for repetition rate 6.9 [F(1, 58)= 13.665, p<0.05], 10.9 [F(1, 58)= 11.048, p<0.05] , 15.4 [F(1, 58)= 5.069, p< 0.05]. MANOVA results showed significant difference in F2 amplitude across the groups for repetition rate 6.9 [F(1, 58)= 12.999, p<0.05], 10.9 [F(1, 58)= 10.323, p<0.05] , but no significant difference for 15.4 [F(1, 58)= 2.023, p> 0.05].

Pearson Correlation between Age and Sustained Components:

In MANOVA, after covarying it with respect to wave V, it was found that for wave D, there was no significant difference across the groups, whereas a significant difference was found for peak E & F. So, to understand the significant difference, a correlational analysis was done where age was the independent variable and wave D, E & F were the dependent variables. Correlation analysis reveals the following results in figure 8 in terms of scatter plot.

Discussion

The present study was conducted with an aim of studying the brainstem correlates of the auditory temporal processing in the young age and older adult adults with normal hearing sensitivity. This was done by recording speech and click evoked ABR at different repetition rates. The two stimuli were chosen as they differ significantly in their acoustic properties.

Effect of Repetition Rate on Latency of Onset Response of Click ABR and Speech Evoked ABR

Present study supports the earlier studies which report a prolongation in latency with increase in repetition rate (Thornton & Coleman 1975; Don, Allen & Starr, 1977; Yagi & Kaga 1979; Lasky, 1984, 1997; Burkard & Hecox 1983, 1987a, 1987b). Burkard and Sims (2001) reported that with increasing click rate, peak latencies increased, the I-V interval increased and peak amplitudes decreased in both young and older normal individuals. Mamatha and Barman (2008) reported that the latencies of wave I, III and V increased with increase in repetition rate within the age groups from 30 to 65 years.

However, there are studies which report that there is no change in latency of click evoked ABR with increase in repetition rate up to 20/sec (Jewett, Romano & Williston, 1970; Krizman, Skoe & Kraus, 2010). Fowler and Noffsinger (1983) also reported no change in latency of click evoked ABR waves with increase in repetition rate between 2- 20 Hz. However, in present study a significant difference was obtained between 9.1/sec and 19.1/sec repetition rates. The difference reported in the present study might be due to the methodological differences between the present and the earlier studies. Earlier studies have been utilized either 80 dB nHL or 90 dB nHL intensity to record auditory brainstem responses whereas, the present study has been done at 80 dB SPL intensity.

With respect to the onset response of the speech evoked ABR, several authors have reported an increase in the onset response with the increase in repetition rate of the stimuli in adults (Krizman, Skoe & Kraus, 2010) and in children (Ranjan, 2011, Mehta & Singh, 2012). The increase in latency of wave V of click ABR and wave V and A of speech evoked ABR due to increase in the repetition rate might be due to cumulative neural fatigue and adaptation, and incomplete recovery involving hair-cell-cochlear nerve junction and also subsequent synaptic transmission. Latency shifts seen with increase in rate in normals may also be due to a change in cochlear receptor functions (Don et al., 1977), the refractory period of individual nerve fibers resulting in a desynchronization of the response that most affects the encoding of the faster elements of the stimulus (Hall, 1992; Jacobson, Murray & Deppe, 1987), decrease in synaptic efficiency (Pratt & Sohmer, 1976) due to which conduction rate decreases and there is an increase in latency. The effect of rate would be additive as the synapses increases from wave I to wave V (Hall, 1992).

The present study also revealed that the latency of the onset responses was more for 40-55 years age group compared to the 18 to 30 years age group for a higher repetition rate and even at lower repetition rates also. Mamatha & Barman (2008) reported that the latencies of wave I, III and V increased with increase in repetition rate across the age groups from 30 to 65 years and there was a greater increase in the latency for wave III and wave V in older individuals. Patterson, Michalewski, Thompson, Bowmanm and Litzelman (1981) reported that older elderly individuals (60 to 79 years) had longer latencies at wave III and wave V compared to the middle aged individuals (40 to 59 years) and middle aged individuals had longer latencies compared to the young adults (20 to 39 years).

These delayed latencies in the onset of the click and speech evoked ABR with increasing age could be consistent with a reduction in synchronous neural firing to transient changes in stimulus and impaired neural encoding of the onset of a stimulus in the older adult individuals. Akhoun et al. (2008) suggested that the onset response of the ABR particularly reflects the synchronous response of many types of brain stem cells at the levels of the cochlear nucleus and inferior colliculus. Therefore, this portion of the response is likely to be affected by age-related loss in neural synchrony in the central auditory system, which may be independent of changes at the periphery (Boettcher, Mills, Swerdloff & Holley, 1996; Gates, Feeney & Higdon, 2003; Mills, Schmeidt, Schulte & Dubno, 2006; Pichora-Fuller, Schneider & McDonald, 2007). This provides an index for examining the role of subcortical timing and its relationship to normal, impaired and the expert auditory perception.

Further, as the repetition rate increased, the difference was maintained between the two groups i.e. the latency was consistently more for the older adult group compared to the younger adults even at higher repetition rate. The neurophysiological mechanisms responsible for observed latency shifts at higher repetition rates in the older adults might be due to taxing the auditory system at higher repetition rates resulting in cumulative neural fatigue and adaptation, and incomplete recovery involving hair-cell-cochlear nerve junction and also subsequent synaptic transmission. This phenomenon might be affected more in older adult individuals compared to the young adults probably because of reduced neural synchrony in older adult individuals. The findings observed at higher repetition rates also suggest an impaired temporal processing in older adults. Behavioral studies have also reported that temporal processing is affected in older adult individuals (Babkoff, Ben-

Artzi & Fostick, 2011; Abel, Krever & Alberti, 1990; Grose, Hall & Buss, 2006). Impaired temporal processing in older adults might be due to reduced neural synchrony, slowed neural conduction time, and reduced phase-locking abilities, which might affect the neurons in the central auditory system to accurately encode important temporal features of signal.

These findings suggest that older adults had a general reduction in synchronous neural firing in response to transient information at the onset of a speech and a click stimulus. Thus, one can hypothesize that the degradation in the onset response of the auditory brainstem responses might start in the older adults itself.

Effect of Repetition Rate on Latency of Speech evoked FFR-Sustained Measure

As the repetition rate increased, a significant prolongation in the latency of few peaks of sustained responses was obtained for both the young adults as well as the older adult group. Looking at the prolongation of the wave V latency of speech evoked ABR, it was suspected that the latency prolongation of the sustained responses might be due to the prolongation of wave V with increase in repetition rate. Hence, the peaks of sustained response were covaried with the wave V latency and after covarying, the rate effect disappeared for all the peaks of older adult group suggesting that the shift seen at sustained responses were a carryover of the large effect of rate on wave V latency for the older adult group. However, after covarying the latencies for the young group, there was a significant difference for wave D and wave F but not for wave E. Krizman et al. (2010) have reported no effect of repetition rate on the sustained responses of the speech evoked ABR in younger adults. The present study followed the recording protocol of Krizman et al. (2010). Although the recording protocol was same, the results obtained were different for the two studies. At this point of time it is difficult to define why repetition rate selectively affected the latencies of two peaks of the sustained responses for the young group.

After covarying the peaks of sustained responses, a significant difference was obtained for E peak between the two groups at all the repetition rates whereas, a significant difference was obtained for the F peak at 10.9 repetition rate and 15.4 repetition rate but not the 6.9 repetition rate and for peak D, there was no difference obtained between the two groups for any of the repetition rate. To understand this, a correlation analysis was done which revealed that there was no correlation between age and wave D latency (i.e. as the age increased, there was no prolongation in the D peak latency), whereas it revealed a significant correlation for wave E for all the repetition rates and for wave F for 10.9 and 15.4 repetition rate. The differences between the two groups for these peaks suggest a selective prolongation of the wave E and F component of the sustained responses for the older adults.

Clinard, Tremblay and Krishnan (2010) also found some significant age effects for the sustained portion of the S-ABR. Significant correlations with advancing age were reported for latencies of the sustained responses for older adults in the age range of 22- 77 years old. Vander, Kathy, Burns and Kristen (2011) also reported a similar finding for individuals in the age range of 61-78 years. The results of the Clinard et al. (2010) and Vander et al. (2011) are in good general agreement with those of this study.

The neurophysiological mechanism behind the encoding of sustained FFR response is dependent on the integrity of phase locked neural activity in the auditory brainstem (Worden & Marsh, 1968). For the encoding of sustained components (wave E and F), there is a significant difference between the young and older adult group, which suggests a delay in encoding of these components at lower repetition rate for the older adults compared to the young adults at the upper brainstem. The effect continues even at higher repetition rates which suggest a possible reduction in temporal processing in the older adults at the upper brainstem level. Temporal processing is dependent on the neural detection of time-varying acoustic cues which might be affected in older adults as a result of poor neural synchrony (Frisina & Frisina, 1997; Schneider & Pichora-Fuller, 2001).

This effect could result from a reduced neural synchrony in peripheral and/or central auditory changes with age. This delay in latencies reflects disrupted neural synchrony, which may also be related to age-related changes in physiology such as metabolic activity in the cochlea (Mills et al., 2006), levels of inhibitory neurotransmitters (Caspary, Schatteman & Hughes, 2005), or decreased cell counts in auditory nuclei (Frisina & Walton, 2006). Age-related related changes to the capacitance and input resistance of inner hair cells (IHCs) or changes in synapses between IHCs and auditory nerve fibers could also influence the coding of the sustained responses (Moser, Neef & Khimich, 2006). For example, deficits of temporal processing have been found in a group of 40-55 year-old individuals (Grose et al., 2006), reduced gap detection ability in middle-aged women (Helfer & Vargo, 2006), reduced DPOAE amplitudes in normal-hearing middle-aged adults (Dorn et al., 1998), subtle differences in auditory perception between younger and older adult subjects in auditory event-related potential (Alain, McDonald, Ostroff & Schneider, 2004; Geal-Dor, Goldstein, Kamenir & Babkoff, 2006), processing of inter aural phase differences both in behavioral and physiological tasks (Ross, Fujioka, Tremblay & Picton, 2007) demonstrating that age-dependent subtle auditory changes may begin in older adult individuals. Thus the poorer encoding of periodicity at the brainstem level in terms of FFR suggests that age-related decline tends to start in the mid age itself.

Effect of Repetition Rate on Representation of F0, F1 & F2-Pitch and Harmonic Measure

Krizman et al. (2010) reported a significant rate effect on the higher harmonics and not on the coding of F0 and F1. However, in the present study there was a significant effect of repetition rate on encoding of F0, F1 and F2 for younger adults and encoding of F2 in older adults. The results obtained in the present study for the older adult group is similar to Krizman et al. (2010). One thing to be noticed here is that even the repetition rate had greater effect on the latencies of sustained responses for the young group compared to the older adult group. After covarying, in the older adult group, there was no effect of repetition rate on latencies of the sustained responses (wave D, E and F). Since the encoding of the F0, F1 and F2 is dependent upon the sustained responses, the responses obtained here for F0, F1 and F2 might be somehow correlated with the latency of the sustained responses in the young group and older adult group. But this mechanism needs to be further checked with more investigations.

A significant difference was seen in the amplitude of F0 and F2 across the groups for 6.9/sec & 10.9/sec and for all repetition rates in the amplitude of F1. These findings suggest that in the older adult individuals there might be a problem in encoding of these key elements of speech. Vander et al. (2011) also reported reduced phase-locking to the fundamental and harmonic frequency components of speech, as measured by the reduced spectral amplitude for F0, F1, and F2 for individuals in the age range of 22-77 years old. Clinard et al. (2010) also reported a reduction in amplitude of F0 in older individuals.

Reduced encoding of F0, F1 and F2 in older adults is consistent with the interpretation of an age-related decline in phase-locking ability involving the brainstem. However, for the F0 and F1 at 15.4 repetitions rate there was no significant difference obtained between the two groups. This might be due to higher standard deviations recorded for these two components at higher repetition rates. Speech recognition abilities were not assessed in this study; therefore, it is not known whether the age-related differences in coding of F0, F1 and F2 directly relate to difficulty in understanding speech in older adults. It will be of interest to see whether the encoding of F0, F1 and F2 such as those observed in the older adult subjects in this study are correlated with reduced speech perception with and without noise condition. However, in the present study, a relation between the temporal processing abilities in this population and encoding of F0, F1 and F2 was obtained, which suggests that the temporal processing might be affected in the older adult individuals itself.

Conclusions

The increase in latency of speech evoked ABR and click evoked ABR for the older adults suggest that the brainstem timing might be affected for the older adults. Both transient and sustained responses of speech evoked ABR shows a significant difference between the young and the older adults suggesting that both the transient and sustained responses are important while doing speech evoked ABR. The peripheral hearing sensitivity was intact in both the groups considered for the study, but there was a reduction in amplitude for the coding of F0, F1 and F2 for the older adult group. Reduction in coding of F0, F1 and F2 might be leading to the speech perception problems in older individuals. Although the perception of speech requires lot more component, brainstem coding of speech sounds might be one of the neural code which might be leading to the speech perception problems in older adults. The study can be utilized to study the subcortical coding of speech at the brainstem level in younger and the older adults. This knowledge could lead to objective diagnostic tests as well as techniques to determine appropriate intervention strategies and ways to monitor the effectiveness of intervention in the elderly population. The data obtained helps us to understand how the temporal aspect of speech and non speech sound is coded at the brainstem level. It highlights the necessity of further studies in different clinical population.

References

Abel, S. M., Krever, E. M., & Alberti, P. W. (1990). Auditory detection, discrimination and speech processing in ageing, noise-sensitive and hearing-impaired listeners. *Scandinavian Audiology, 19,* 43-54.

Akhoun, I., Gallego, S., Moulin, A., Menard, M., Veuillet, E., Berger-Vachon, C., et al. (2008). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clinical Neurophysiology, 119,* 922-933.

Alain, C., McDonald, K. L., Ostroff, J. M., & Schneider, B. (2004). Aging: a switch from automatic to controlled processing of sounds? *Psychology and Aging, 19,* 125-133.

American National Standards Institute. (1991). *American National Standard Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms*. ANSI S3.1- (1991). New York: American National Standards Institute.

Babkoff, H., Ben-Artzi, E., & Fostick, L. (2011). Auditory temporal processes in the elderly. *Audiology Research, 1(6),* 19-21.

Barr, R. A., & Giambra, L. M. (1990). Age-related decrement in auditory selective attention. *Psy-*

chology and Aging, 5, 597-599.

Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Developmental Science, 13,* 77- 91.

Bergman, M. (1971). Changes in hearing with age. *Gerontologist, 11,* 148-151.

Boettcher, F. A., Mills, J. H., Swerdloff, J. L., & Holley, L. (1996). Auditory evoked potentials in aged gerbils: Responses elicited by noises separated by a silent gap. *Hearing Research, 102,* 167-178.

Burkard, R., & Hecox, K. E. (1983). The effect of broadband noise on the human brainstem auditory evoked response. I. Rate and intensity effects. *Journal of the Acoustical Society of America, 74,* 1204- 1213.

Burkard, R., & Hecox, K. E. (1987a). The effect of broadband noise on the human brainstem auditory evoked response. III. Anatomic locus. *Journal of the Acoustical Society of America, 81,* 1050- 1063.

Burkard, R., & Hecox, K. E. (1987b). The effect of broadband noise on the human brainstem auditory evoked response. IV. Addidivity of forward- masking and rate- induced wave V latency shifts. *Journal of the Acoustical Society of America, 81,* 1064- 1072.

Burkard, R., & Sims, D. (2001). The human auditory brainstem response to high click rates: Aging effects. *American Journal of Audiology, 10*(2), 53-61.

Caspary, D. M., Schatteman, T. A., & Hughes, L. F. (2005) Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: role of inhibitory inputs. *Journal of Neuroscience, 25,* 10952-10959.

Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264,* 48 -55.

Committee on Hearing and Bioacoustics and Biomechanics (CHABA)(1988). Speech understanding and aging. *Journal of the Acoustical Society of America, 83,* 859-895.

Don, M., Allen, A. R., & Starr, A. (1977). Effect of click rate on the latency of auditory brain stem responses in humans. *Annals of Otology, Rhinology & Laryngology, 86,* 186- 195.

Dorn, P. A., Piskorski, P., Keefe, D. H., Neely, S. T., & Gorga, M. P. (1998). On the existence of an age/threshold/frequency interaction in disortion product otoacoustic emissions. *Journal of the Acoustical Society of America, 104,* 964-971.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *Journal of the Acoustical Society of America, 111,* 2897-2907.

Era, P., Jokela, J., Qvarnberg, Y., & Heikkinen, E. (1986). Pure-tone thresholds, speech understanding, and their correlates in samples of men of different ages. *Audiology, 25,* 338-352.

Ewertsen, H. W., & Birk-Nielson, H. B. (1971). A comparative analysis of the audiovisual, auditive and visual perception of speech. *Acta Oto-Laryngologica, 72,* 201-205.

Fitzgibbons, P. J., & Gordon-Salant, S. (1996). Auditory temporal processing in elderly listeners, *Journal of American Academy of Audiology, 7,* 183-189. Fowler, C. G., & Noffsinger, D. (1983). Effects of stimulus repetition rate and frequency on the auditory brainstem response in normal cochlear-impaired, and VIII nerve/brainstem- impaired subjects. *Journal of Speech and Hearing Research, 26,* 560- 567.

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research, 106*, 95-104.

Frisina, R. D., & Walton, J. P. (2006). Age-related structural and functional changes in the cochlear nucleus. *Hearing Research, 216-217*, 216-223.

Gates, G. A., Feeney, M. P., & Higdon, R. J. (2003). Word recognition and the articulation index in older listeners with probable age-related auditory neuropathy. *Journal of the American Academy of Audiology, 14,* 574 -581.

Geal-Dor, M., Goldstein, A., Kamenir, Y., & Babkoff, H. (2006). The effect of aging on event-related potentials and behavioral responses: comparison of tonal, phonologic and semantic targets. *Clinical Neurophysiology, 117,* 1974-1989.

Gelfand, S. A., Piper, N., & Silman, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *Journal of the Acoustical Society of America, 80,* 1589-1598.

Gordon-Salant, S. (1987). Age-related differences in speech recognition performance as a function of test format and paradigm. *Ear and Hearing, 8,* 270-27.

Grose, J. H., Hall, J. W., & Buss, E. (2006). Temporal processing deficits in the pre- senescent auditory system. *Journal of the Acoustical Society of America, 119,* 2305-2315.

Hall, J.W. (1992). *Handbook of Auditory Evoked Responses*. Boston: Allyn & Bacon.

Helfer, K., & Vargo, M.(2006). Speech Recognition and Temporal Processing in Middle-Aged Women. *Journal of the American Academy of Audiology, 20,*(4) 264-271.

Humes, L. E. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology, 7,* 161-167.

Jacobson, J. T., Murray, T. J., & Deppe, U. (1987). The effects of ABR stimulus repetition rate in multiple sclerosis. *Ear & Hearing, 8,* 115-120.

Jewett, D. L., Romano, M. N., & Williston, J. S. (1970). Human auditory evoked potentials: Possible brain stem components detected on the scalp. *Science, 167,* 1517-1518.

Kim, S., Frisina, R. D., Mapes, F. M., Hickman, E. D., & Frisina, D. R.(2006). Effects of age on binaural speech intelligibility in normal hearing adults. *Speech Communication, 48,* 591-597.

Klatt, D. (1980). Software for cascade / parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971- 975.

Krizman, J., Skoe, E., & Kraus, N. (2010). Stimulus Rate and Subcortical Auditory Processing of Speech. *Audiology Neurotolgy, 15,* 332-342.

Lasky, R. E. (1997). Rate and adaptation effects on the auditory evoked brainstem response in human newborn and adults. *Hearing Research, 111,* 165- 176.

Mamatha, N. M. & Barman, A. (2008). Effect of repetition rate on auditory brainstem response in adults and elderly. *Articles based on dissertation at A.I.I.S.H, 1,* 39-51.

Mazelova, J., Popelar, J. & Syka, J. (2003). Auditory function in presbycusis: peripheral vs. central changes. *Experimental Gerontology, 38,* 87-94.

Mehta, G., & Singh, N. K. (2012). *Stimulus rate and subcortical auditory processing of speech in children with learning disability.* Poster presented at 44th ISHACON, Hyderabad.

Mills, J. H., Schmeidt, R. A., & Schulte, B. A., & Dubno, J. R. (2006). Age-related hearing loss: A loss of voltage, not hair cells. *Seminars in Hearing, 27,* 228 -236. Moser, T., Neef, A., & Khimich, D. (2006). Mechanisms underlying the temporal precision of sound coding at the inner hair cell ribbon synapse. *Journal of Physiology, 576*(1), 55-62.

Nabelek, A. K., & Robinson, P. K. (1982). Monaural and binaural speech perception in reverberation for listeners of various ages.*Journal of the Acoustical Society of America, 71,* 1242-1248.

Patterson, J. V., Michalewski, H. J., Thompson, L. W., Bowmanm, T. E., & Litzelman, D. K. (1981). Age and Sex Differences in the Human Auditory Brainstem Response. *Journal of Gerontology, 36*(4), 455-462.

Pichora-Fuller, M. K., Schneider, B. A., & Mcdonald, E., et al. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research, 223,* 114 -121.

Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *Journal of the Acoustical Society of America, 66,* 1333-1342.

Pratt, H., & Sohmer, H. (1976) Intensity and rate functions of cochlea and brainstem evoked responses to click stimuli in man. *Archives of Otorhinolaryngology, 212,* 85-93.

Ranjan, R. (2011). Effect of stimulus rate on Subcortical auditory processing in children. Unpublished dissertation done at A.I.I.S.H, Mysore. Ross, B., Fujioka, T., Tremblay, K. L., & Picton, T. W. (2007). Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *Journal of Neuroscience, 27,* 11172-11178. Schneider, B. A., & Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: implications for speech perception. *Seminars in Hearing, 22,* 227-39.

Thornton, A. R. D., & Coleman, M. J. (1975). The adaptation of cochlear and brainstem auditory evoked potentials in humans. *Electroencephalography and Clinical Neurophysiology, 39,* 399- 406.

Vander, W., Kathy, R., Burns, & Kristen, S. (2011). Brainstem responses to speech in younger and older adults. *Ear & Hearing, 32(2),* 168-180. Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., & Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *Journal of the American Academy of Audiology, 17,* 487-497.

Worden, F. G., & Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology, 25, 42-52.*

Yagi, T., & Kaga, K. (1979). The effect of the click repetition rate on the latency of the auditory evoked brainstem response and its clinical use for a neurological diagnosis. *Archieves of Otorhinolaryngology, 222,* 91- 97.

Effect of Hearing Aid Channels on Acoustic Change Complex

[1]Giten Eliza George & [2]Sreeraj Konadath

Abstract

The objective of the present study was to find the effect of varying the number of channels of a hearing aid on the Acoustic change complex (ACC) in individuals with sloping hearing loss and to verify the same using the speech identification scores. The ACC was obtained from individuals with normal hearing and those with sloping sensorineural hearing loss in response to /si/ stimulus, the comparison in ACC wave forms were then made between these two groups. The ACC obtained across two, four and eight channel hearing aids with the same make and manufacturer were compared. A high frequency content speech identification test was done as a behavioral correlation measure. The results showed that ACC latencies in response to the fricative portion /s/ were delayed in individuals with sloping hearing loss. Between the two, four and eight channels, there were no significant differences found in terms of latency, amplitude, morphology or speech identification scores. The electrophysiological findings were found to poorly correlate with the behavioral speech identification measure. Therefore it was concluded that naturally produced speech tokens, representing different acoustic cues, like frication and vowel steady states can evoke distinct neural response patterns, and cortical evoked potentials elicited by /si/ stimulus can be reliably recorded in individuals with and without hearing aids. There was no significant difference in perception found between the two, four and eight channels.

Keywords: *Acoustic Change Complex (ACC), Cortical Auditory Evoked Potentials (CAEP), Hearing aids, Channels, Speech Perception.*

Introduction

The acoustic change complex (ACC) is a cortical auditory evoked potential that can be obtained from the auditory system when a time varying acoustic change occurs within the stimulus. The change can be in amplitude, spectral envelope or periodicity. Speech is one such stimulus which has multiple time varying acoustic changes, and can be used to elicit this potential (Martin & Boothroyd, 2000). The ACC being an obligatory potential depends solely on the acoustic features of the stimulus and the integrity of the central auditory system, further it does not require active involvement from the participants.

There is interest in using this potential as it can probe into supra-threshold auditory skills, such as speech sound processing. This potential may be able to dwell deeper into the perception of speech as an index of speech sound discrimination as it comprises not only a detection waveform but also one which arises due to a change in the acoustic characteristics of the signal.

Aided late latency responses can be reliably recorded according to Tremblay, Billings, Friesen, and Souza (2006), which will yield a better understanding of how device settings will affect the neural response patterns which in turn affects speech perception. After varying a device settings or a feature, if a cortical response will still mimic stimulus acoustic features, then a conclusion can be drawn about both, speech perception and how the hearing aid processes the signal.

In a sloping hearing loss, as the frequency increases the degree of hearing loss also increases (Roeser & Clark, 2007). Persons with sloping hearing loss due to a cochlear pathology can face problems while listening due to decreased audibility and dynamic range, also secondary to reduced frequency and temporal resolution. While studying the distribution of hearing loss characteristics in terms of configuration, Margolis and Saly (2008), found that sloping hearing losses dominated the distributions of configuration and that sensori-neural was the most prevalent site of lesion. Demeester et al. (2008) found the prevalence of high-frequency gently sloping and steeply sloping hearing loss to be 76% in males and 50% in females between 55 to 65 years. This means that majority of the persons with hearing loss miss the high frequency portions of speech, which are consonants. Consonants as compared to vowels were found to be approximately 30 dB lower in intensity and are hence less audible (Zeng & Turner, 1990). Hearing aids can help alleviate these issues by selective amplification, as in by increasing only the gain of higher frequency components.

Selective amplification can be accomplished through hearing aids by various means, one of which is the use of multichannel hearing aids which split the incoming signal into frequency bands. A band in a hearing aid refers to a frequency region where gain adjustments are made and a channel is where the same signal processing takes place. Individual compression circuits allow these multichannel instruments the flexibility to amplify each bandwidth of frequencies independently so as to correspond to the user's needs, preferences and their dynamic range. In addition, each channel may be set with unique

[1]Email: gitenegeorge@gmail.com,
[2]Lecturer in Audiology, Email: sreerajkonadath@aiishmysore.in

attack and release time for compression across the frequency range as reported by Dillon (2001).

The number of channels necessary for optimal speech perception varies across studies. There are various factors involved that can dictate the amount of benefit a multichannel hearing aid can provide. For example, the configuration, age, degree of hearing loss, ability to combine temporal-envelope information and hearing aid experience (Jyoti, 2010; Rubina, 2008; Souza & Boike, 2006; Yund & Buckles, 1995). There are conflicting finding about the benefit derived from multi-channel hearing aids as authors have found that these hearing aids may reduce the spectral contrast which aids in the perception of vowels (Bor, Souza & Wright, 2008).

According to Plomp (1988), this occurs when there are multiple channels with large compression ratios. However, Yund and Buckles (1995), demonstrated improvement in speech recognition from four to eight channels. Jyoti (2010) compared two, four and eight channels in listeners with sloping hearing loss and reported of better speech identification scores with increased number of channels. As there are contrasting results, electrophysiological verification of the benefits or detrimental effects of having multiple channels has been explored and behavioral speech identification scores were also obtained. Therefore, conclusions can be drawn on the effect of multi-channel hearing aids on the central auditory system processing of speech.

The P1-N1-P2 neural response patterns are heavily influenced by the acoustic content of the evoking signal and hearing aids may alter the acoustic content which will degrade speech perception. For example, the perception through hearing aids can blur the boundary between the aperiodic noise of consonants and the onset of voiced vowels, making these transitions less distinct (Stelmachowicz, Kopun, Mace, Lewis & Nittrouer, 1995).

To understand the interaction between the digitally amplified signal and its neural representation, physiologic detection of CV transitions was studied in a group of individuals who were first time users of hearing aids. The results of this study can be used as a counseling tool which can help clients make informed decisions and understand why higher or lower number of channels are required to suit their need.

The objectives of this study were to, one, compare ACC in individuals with normal hearing sensitivity to those with sloping cochlear hearing loss. Second, to compare the aided and unaided ACC in individuals with sloping cochlear hearing loss. Third, to compare the performance across two, four and eight channels of hearing aid. Fourth, to correlate speech identification score and the ACC obtained across different number of channels.

Method

Participants

Two groups, the control and the clinical group were included in the study. In the control group16 individuals (20 ears; equal number of male and female ears) in the age range of 25 to59 years were included. The criteria for inclusion comprised normal hearing sensitivity (air conduction and bone conduction thresholds less than 20 dB HL across frequency range from 250 Hz to 8 kHz) in both ears with normal middle ear status (defined as peak pressure between -100 and +60 daPa, and admittance between 0.3 and 1.60 cc). They were also required to be native Kannada speakers.

In the clinical group, 10 native Kannada speakers (11 ears; 3 females and 7 males) in the age range of 25 to59 years were included. The criteria for selection included air conduction thresholds which increased by 5 to 12 dB per octave and speech identification score of >55% in the test ear.

Instrumentation and Test Environment

A calibrated double channel audiometer, GSI- 61 was used to estimate the pure tone thresholds and to obtain speech identification scores. The GSI Tympstar (version-2) middle ear analyzer was used for immittance measurements. For electrophysiological recording a 2-channel diagnostic auditory evoked potential measuring instrument, Bio-Logic Navigator Pro (version 7.0.0) was used to record the ACC waveform.

To present the speech stimulus for ACC in aided and unaided condition a calibrated dB Technologies-160 free field speaker was used which had a frequency response range of 50 Hz to 20,000 Hz and maximum sound pressure level of 99dB SPL.

Fonix 7000 hearing aid test system was used for electroacoustic measurement of hearing aids and for real ear measurement.

The recording of the test stimuli and the audiological testing were done in acoustically treated rooms, with the noise levels within permissible limits according to ANSI S3.1, 1991. Pure tone audiometry was done in a double room suite, and the ACC acquisition in a single room set-up.

Test Procedure

The procedure started with taking a detailed case history probing into any history of ear related pathologies. Pure tone thresholds were obtained in octave intervals between 250 Hz to 8000 Hz for air conduction and between 250 Hz and 4000 Hz for bone conduction using the modified Hughson-Westlake procedure (Carhart & Jerger, 1959). Tympanometry and reflexometry were

Table 1: Stimulus and Acquisition parameters of ACC

Stimulus parameters		Acquisition parameters	
Stimuli	/si/	Mode of stimulation	Ipsi
Duration of stimulus	250 ms	Electrode montage	Cz, M1/M2 of test ear, ground at Fz
Intensity	65 dB SPL		
Polarity	Alternating	Filter setting	0.1-30 Hz.
Transducer	Loudspeaker	Analysis window	535 ms.
Mode of presentation	Free Field	No. of channels	Single
		Amplification	25,000
		Repetition rate	1.1 per second
		Number of sweeps	150
		No. of repetitions	2

Table 2: Acoustic features of the /si/ stimulus

Acoustic Features	Values
Total Duration	250 ms
Fricative Duration	143 ms
Vowel Duration	107 ms
Fricative Center Frequency	Energy spread from 2 to 5 kHz
Fundamental frequency - Vowel	75 Hz
First Formant - Vowel	1059 Hz
Second Formant - Vowel	2557 Hz

done to exclude individuals with middle ear pathology. To accomplish the same, a group of 50 individuals were administered the diagnostic protocol mentioned above, after which 16 of them (20 ears) were included in the study. The method involved 4 phases.

Phase 1: Acquisition of ACC in Control Group

The neural representation of a fricative /s/, its transition to a vowel, and perception of the vowel /i/ can be studied through the ACC. If these elements are perceived at the cortical level, two distinct waveforms will be observed, one in response to the consonant and the other for the vowel.

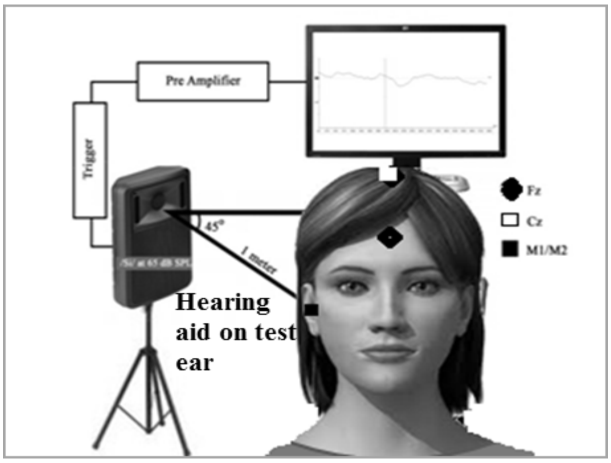


Figure 2: Set-up for aided ACC measurement - Electrode and speaker placement.

To learn the effect of sloping sensorineural hearing loss, the clinical group's unaided ACC waveforms were compared to that of the control group; the differences in latency, amplitude and morphology between the two responses gave information on the effect of sloping sensori-neural hearing loss on ACC. The inclusion of a control group can prevent extrinsic variables from influencing results, as both groups comprised native Kannada speakers, with no experience in listening through an amplification device. Additionally, cognitive abili-

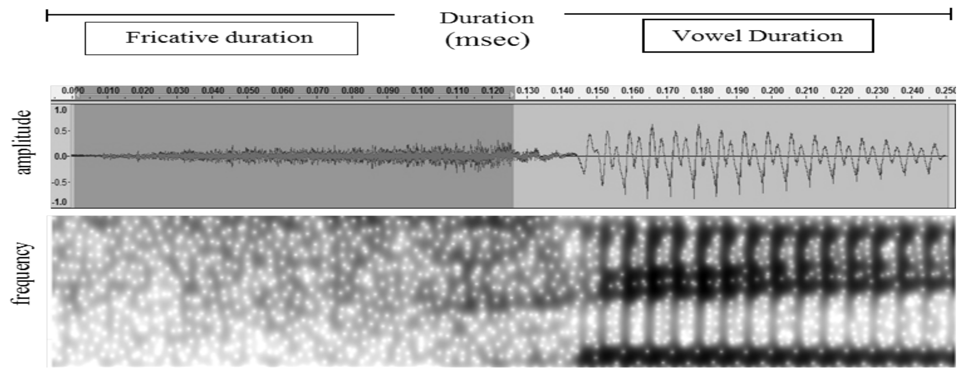


Figure 1: Wave form (top) and spectrogram(bottom) of stimulus /si/.

Table 3: Electro-acoustic measurement of the three hearing aids

Electro-acoustic measure	2 Channel	4 Channel	8 channel
Output Sound Pressure Level (OSPL) at 2100 Hz	130.7 dB	129.8 dB	131.3 dB
HFA- FOG	53.2 dB	49.8 dB	59.3 dB
Reference Test Gain	47.1 dB	46.4 dB	46.3 dB
Frequency Range	200 - 5785 Hz	200 - 5787 Hz	200 - 6212 Hz
EIN	12. 4 dB	10.1 dB	16.0 dB
THD			
500 Hz	2.0%	1.2%	0.5%
800 Hz	1.4%	1.3%	0.4%
1600 Hz	0.1%	0.1%	0.2%

Note: THD = Total Harmonic Distortion, EIN = Equivalent Input Noise, HFA-FOG = High Frequency Average-Full OnGain

ties that might influence the processing speech stimulus at the level of the auditory cortex can be taken into consideration. The stimulus and the acquisition parameters are detailed in Table 1.

While acquiring the ACC, participants were seated comfortably on a reclining chair. The electrode site was cleaned and gold disc electrodes were placed at the test site. The inter electrode impedance was maintained to be less than 2 kΩ and was monitored throughout the recording. The participants were instructed to ignore the stimulus and minimize head movements. A total of 300 sweeps were acquired for each participant. If the waveform was not found to replicate, a third recording was done.

*Stimulus Preparation:*To elicit ACC, /si/ stimulus was naturally produced and recorded using Adobe Audition (version 1.5) at a sampling rate of 48,000 with 16 bit resolution. A dB Technologies-160 free field speaker was used to present the /si/ stimulus for recording ACC for both the aided and unaided conditions. The output of the speaker was calibrated using a Larson-Davis System-824 sound level meter to be presented at a level of 65 dB SPL. The speaker was positioned at a distance of one meter at 45^o azimuth. With the PrattC software (version 5.1.31) the acoustic features of the /si/ stimulus was analyzed and is shown in Table 2, and the stimulus waveform and spectrogram are shown in Figure 1 and Figure 2 respectively.

Phase 2: Hearing aid fitting for the clinical group

Electro-acoustic Measurement: Two, four and eight channel hearing aids of the same make and manufacturer were utilized for the study. The electro-acoustic measurement with the hearing aid in test setting [ANSI S3.22-2003; (2 cc coupler)] was performed as shown in Table 3.

Hearing aid Programming: The NOAH 3 software using the Connexx platform, connected to HiPro was used for programming the hearing aid. The client would be seated with the probe microphone and the programming system placed in the same room to simultaneously match the target while programming the hearing aid. The hearing aids were in omni-directional mode with enabled compression circuits, and compression ratio set according to NAL-NL1. At an input of 65 dB SPL, gain provided across 250 Hz to 6500 Hz was verified. Of the four programmable memories in each hearing aid, only one that is the speech in quiet mode was activated, also the volume control was disabled.

NAL-NL1 fitting formula was used to prescribe the gain across different channels and accordingly the hearing aid was programmed to match the target gain at acclimatization level two. To specify the amplification targets for speech and the maximum output necessary to provide loudness comfort, audibility of speech, and speech intelligibility, this prescriptive formula was used.

To quantify the stimulus level at the output of the hearing aid, a probe-tube microphone system was placed in the ear canal, and sound level measurements were made with the hearing aid. Computation of hearing aid delay was done to remove the effect of the hearing aid processing delay while analyzing the latency. Processing delay of the 2 - channel hearing aid was 1.5 ms, and the 4 and 8 channel aids had 1.8 ms and 3.3 ms delays respectively. However, as the delays were very small compared to the standard deviation of the ACC, it was not taken into consideration during final analysis.

Phase 3: Acquisition of aided ACC for clinical group

The participant was comfortably seated on a recliner in the sound treated room. The electrode placement was as mentioned in Table 1 and depicted in Figure 3. Instrumentation in the single sound treated room used is shown in Figure4. Aided ACC waveforms were acquired, with the hearing aid placed on the test ear. The behind-the-ear hearing aid had to be placed in such a manner, that it caused no interference or artifacts as the inverting electrode was also placed on the test ear mastoid. Next, the speaker was positioned so that the hearing aid microphone, relative to the speaker was at an angle of 45^o. A pre-stimulus electroencephalographic recording was first observed, and once the trace was found to stabilize the test recording was initialized. To obtain a baseline of the neurological activity unrelated to the stimulus in the waveform, a prestimulus recording was done for 100 ms. The waveforms were then analysed.

Analysis of the waveforms: The three-aided ACC waveforms (ACC obtained from 2, 4 and 8 channel) were compared based on their respective latency, amplitude and morphology. First, the waveforms were visually judged to be replicable after which peak latency and amplitude measures were made. The waveforms were analyzed by two experts in the electrophysiological measurement of auditory evoked potentials. Absolute amplitude and latency were chosen at the most negative or positive point or halfway point of a broad peak was considered.

Secondly, the morphology was rated based on three point rating scale. A score of 0 indicates poor morphology, 1 and 2 indicates average and good wave form morphology respectively. The data was tabulated for N1, P2, P1' and N1'.

Phase 4: Obtaining the speech identification score in the clinical group

Aided speech identification scores were obtained using the High Frequency Kannada Speech Identification Test (Mascarenhas, 2002). Each channel of the hearing aid was evaluated with one of the four lists consisting of 25-words each; an unaided SIS score was also obtained.

This word list had been developed exclusively for high frequency sloping sensorineural hearing loss. In an acoustically shielded room, the voice of a native Kannada female speaker was recorded. A female voice was used as it has a higher fundamental frequency, and can better tap into perception of high frequency phonemes. Speaker effects were eliminated with the use of words which were recorded using Adobe Audition (version 1.5) with 16 bit resolution. To prevent familiarization, three word list of the original test were utilized along with a fourth, that was developed through randomization of the words in the first list. Prior to each list, a 1000 Hz calibration tone was recorded in each word list, and was used to adjust the VU meter of the audiometer to zero. The words were presented at 65 dB SPL through a diagnostic 2-channel audiometer at 0^0azimuth.

The results obtained with the above procedure was subjected to statistical analysis performed using the Statistical Package for Social Sciences (SPSS) (version 16).

Results and Discussion

Comparison of Unaided Scores between Clinical and Control Group

The first objective was to compare the ACC of participants with sloping hearing loss (clinical group) with the participants having hearing sensitivity within normal limits (control group). With a descriptive analysis of the control and clinical group, the mean and the standard deviation values could be obtained and is shown in Table 4. The N1-P2 complex was generated in response to perception of the fricative /s/ and the N1'-P2' complex in response to the vowel /i/.

To study the outcome Mann-Whitney U test was done for comparison of the two groups. The test revealed statistically significant differences between latency of P2 [/Z/ = -02.405, $p < 0.05$] which was significantly shorter in the control group. The amplitude of N1' [/Z/ = -0.377, $p < 0.001$], was significantly larger in the control group. Although the mean latency of the N1' peak for the clinical group is shorter, it was not found to be significant on statistical analysis.

A noticeable feature depicted in Table 4, which compares the mean latencies of the control and clinical group is the absence of the N1 peak in the clinical group participants. The N1 latency and amplitude values could not be considered for comparison as none of the clinical group participants had this peak present in their waveforms. As the N1 reflects detection of the stimuli at the cortical level (Martin, Tremblay & Stapells, 2007), the absence of the same in the sloping cochlear hearing loss individuals may have arose due to three main reasons. The first reason is the inaudibility of the fricative portion of the /si/ stimulus. From a visual analysis of the /s/ portion shown in Figure 1, it is evident that the fricative portion has lower amplitude than the vowel portion. The first N1-P2 complex is generated in response to the fricative portion of the stimulus and the second N1'-P2' complex is generated in response to the transition from fricative to the vowel. The transition portion is audible to those with sloping cochlear hearing loss as evidenced by the presence of the N1'-P2' complex, but the initial /s/ portion is not audible.

The P2 peak has an earlier latency in the control participants when compared to the clinical group, but, the P2 amplitude was similar for both the groups. This could be because the control group participants could perceive the initial fricative /s/, but, due to reduced audibility there was a delay in latency for the clinical group. There were no differences between the two groups in the amplitude of the P2 peak in the present study, and is comparable to results found by Wall, Dalebout, Davidson and Fox (1991). These results have also been confirmed in literature by Oates, Kurtzberg and Stapells (2002),

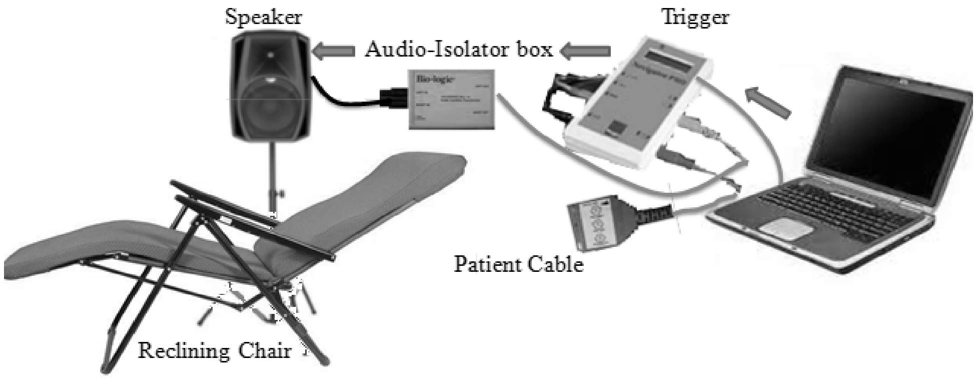


Figure 3: Equipment set-up for ACC measurement.

Table 4: Unaided parameters obtained for the control and clinical groups

Parameter Analyzed	CONTROL			CLINICAL		
	N	Mean	SD	N	Mean	SD
N1 latency (ms)	11	127.05	15.34	-	-	-
N1 amplitude (μV)		-01.35	.925	-	-	-
P2 latency (ms)	20	179.74	26.59	6	201.07	11.68
P2 amplitude (μV)		1.610	.873		1.30	1.05
N1' latency (ms)	20	256.98	11.77	8	244.89	20.50
N1' amplitude (μV)		-2.84	1.20		-2.79	1.73
P2' latency (ms)	20	316.15	20.03	9	317.93	25.93
P2' amplitude (μV)		1.06	1.14		2.67	1.27
Waveform Morphology	20	1.55	0.51	11	1.40	0.843

Note:Lat = Latency (ms), Amp = Amplitude (µv), Morph = Morphology

who found latency measures to be a more sensitive indicator of the effect of decreased audibility than a response strength measure such as amplitude.

Similar findings have been reported by Polen (1984) who found significantly longer latencies for P2 in listeners with hearing loss compared to listeners with normal hearing. The author stated that this could be because this peak is sensitive to stimulus features and a reduction in audibility of these features causes the peak to diminish.

Although the frication portion was audible to the control group, nine of the twenty control group participants did not have the N1 peak in their waveform. Therefore, there could be features inherent to the stimulus

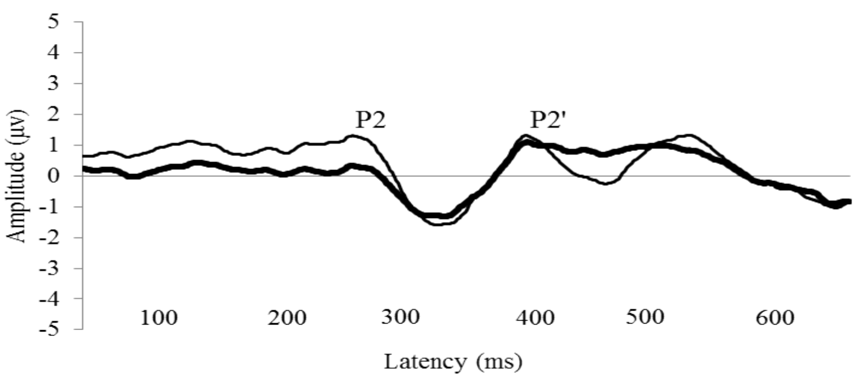


Figure 4: Grand mean ACC waveform of participants with hearing sensitivity within normal limits (thin) and sloping sensorineural hearing loss (thick).

Table 5: Latency, amplitude and morphology (mean and standard deviations for two, four and eight channel hearing aids)

Parameter Analyzed	8-Channel			4-Channel			2-Channel		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
P2 Lat	8	201.72	19.00	8	209.04	13.24	6	190.47	17.21
P2 Amp		1.00	0.67		1.49	0.93		0.84	0.69
N1' Lat	10	251.48	14.62	10	254.63	19.03	9	251.10	17.54
N1' Amp		-2.64	1.49		-2.61	1.10		-2.78	1.50
P2' Lat	10	314.00	20.58	10	317.33	24.20	9	316.92	27.80
P2' Amp		2.30	1.37		1.97	1.44		2.22	1.29
Morph	11	1.45	0.69	11	1.28	0.79	11	0.82	0.75

as in how the vowel and consonant pair interact or the way the central auditory system responds to high frequency stimuli that may have caused the N1 peak to be absent. To explain further, the second reason considers the masking phenomenon; individuals with cochlear pathology are more prone to the effects of upward spread of masking due to wider auditory filters, which in turn results in reduced frequency resolution (Moore, 1998).

The vowel portion of the stimulus could have sufficiently masked the fricative causing a response to arise only from the vowel portion. Low frequency portions of speech like vowels may mask higher frequency components like frication noise (Dillon, 2001). The /i/ portion being larger in amplitude could have masked the /s/ portion of the stimulus used in this study. The third reason is the inherent lower amplitude of the cortical response when elicited from stimulus containing high frequency content such as /ʃ/ and /s/ compared to /m/, /a/, /u/ and /i/ which had predominant low frequency content as documented by Agung, Purdy, McMahon and Newall (2006). The N1 amplitude in their study was significantly smaller for the stimuli containing high frequency content. As responses evoked by a high frequency stimulus have inherent lower amplitude, an additional cochlear pathology could have resulted in an elimination of the response.

The vowel /i/ is capable of eliciting earlier latency. This is because the high front vowel such as /i/ can evoke CAEPs that have earlier latencies compared to low mid-back vowels /u/. Vowels with large F2-F1 differences such as /i/ ($\tilde{2}300$ Hz) and /u/ ($\tilde{1}700$ Hz) have larger areas of activation compared to a vowel with a small F2-F1 distance such as /a/ according to Yetkin, Roland, Christensen and Purdy (2004).

In listeners with normal hearing sensitivity, the growth of loudness at mid and high level follow a compressive power function of intensity (Stevens, 1955). Whereas, in cochlear hearing loss, a loss of compression occurs, this is consistent with the presence of recruitment. According to Buus and Florentine (2001), this growth of loudness occurs near threshold; on average for every 16 dB of hearing loss, the loudness near threshold doubled. The finding of a larger P2' amplitude in the clinical group can be explained based on the differences in the growth of loudness between the control and clinical group. The increase in the amplitude of P2 in the clinical group may be attributed to recruitment which could have occurred as a result of cochlear hearing loss. The grand mean waveforms of the control and the clinical group ACC has been depicted in Figure 5. The amplitude differences have been shown in Table 4. The second objective was to compare the unaided and aided ACC in individuals with sloping cochlear hearing loss. Comparison between the three aided conditions and the unaided condition was done. The aided conditions considered three hearing aids comprising two, four and eight channels.

Comparison of the unaided and aided conditions within the clinical group

Wilcoxon Signed Ranks Test result did not reveal any significant differences between the aided (two, four and eight channels) and unaided conditions ($p > 0.05$). Therefore, for individuals with sloping cochlear hearing loss, the performance on the ACC did not differ with amplification or without it. While studying the neural representation of amplified speech sounds, similar findings were reported by Tremblay et al., (2006). An increase in hearing aid gain did not result in an amplitude increase or latency decrease as compared with unaided ACC at the same input level. They found no significant differences between aided and unaided ACCs in response to the /si/ stimulus. Reasons suggested for the findings have been hearing aid compression effects, which is also applicable in this study as the compression circuit was active during ACC acquisition. To explain further, according to Easwar, Glista, Purcell and Scollie (2012) inter stimulus interval between one and two seconds causes an abrupt increase in the input level each time the stimulus is presented. This causes compression to act each time, and may result in overshoot for every stimulus presented. This overshoot occurs when the stimulus is above compression threshold, because the hearing aid requires time to stabilize the gain when there is a rapid change in the input level.

Another reason could be that the amount of response change (improvements) seen in the ACC demonstrated considerable variability across subjects. This could have led to insignificant findings in the current study. One of the participants in the clinical group had no visible ACC in the unaided condition or in the aided although the stimulus was audible. The finding of this subject was not included for statistical analysis.

Effect of hearing aid channels

Speech cues like transitions with dynamic frequency changes are not perceived well even when audibility is provided. Zeng and Turner (1990) stated the loss of transition cue perception which occurs in sensorineural hearing loss cannot be compensated by hearing aids. In contrast, the hearing aids in the present study could reflect the transition portion as evidenced by the prominent ACC N1'- P2'peak.

Each peak obtained in the ACC, was analyzed for latency, amplitude and morphology. The descriptive statistics has been shown below; in Table 5. To learn the effect of hearing aid channels on the central auditory system, the ACC latency, amplitude and morphology across 2, 4 and 8 channels were compared. The Friedman's test was used and the results obtained have been discussed below. No significant differences were found between the two, four and eight channels in terms of latency and amplitude; this means that the three hearing aids had equivalent performance. The same has been shown in Figure 6 and Figure 7. It has been reported that the vowel-consonant differences will be reduced due to spectral contrast reduction caused by a higher number of channels (Plomp, 1988).

Vowels are more susceptible to this effect than consonants. Consonants in the initial position are more susceptible than those in the final position (Boothroyd, Mulhearn, Gong & Ostroff, 1996). However, the results of this study contradicts the findings of Plomp (1988) and Bor et al. (2008), as the two, four and eight channel hearing aids were not found to have significant differences in latency and amplitude. The effects of spectral contrast reduction which would have reduced the difference between the vowel and the fricative would have otherwise led to increased latency, reduced amplitude and poor morphology for the eight channel hearing aid. Among the four parameters analyzed, the waveform morphology was found to be statistically significant across the three aided conditions with $\chi^2(11) = 7.00$, p = < 0.05. Next a Wilcoxon Signed Rank test was conducted and the results showed that morphology obtained with the four and eight channel devices were equivalent. However, between the eight and two channel hearing aids there was a significant difference in morphology, /Z/ = -02.111, $p < 0.05$. The mean of the ranks in favor of a 2 channel hearing aid were 4.00, while the mean of the ranks in favor of eight channels were 4.57. From the descriptive statistics, mean value of the 8 channel hearing aid was higher than that of the 2 channel hearing aid. Therefore, it can be concluded that an 8 channel hearing aid results in better waveform morphology relative to the 2 channel hearing aid, Figure 8 compares the grand mean waveforms of the three aided conditions. Since poor waveform morphology may not have a physiological basis, and can be affected by a number of factors from sweep to sweep such as muscular artifacts, eye blink and state of arousal, it may not be consistent with changes in signal processing.

Multichannel hearing aids do not have rectangular analysis bands and compression channels, but instead the actual analysis bands and compression channels in hearing aids are non-rectangular, and may have very gradual filter slopes. The more gradual the filter slope of the channels,the compression across channels will be correlated or similar. Therefore, a hearing aid with higher number of channels may be functioning similar to a hearing aid with lower number of channels. This could be a probable reason for why no significant changes were noted across two, four and eight channels.

Another reason for no significant differences across channels could be due to the compression circuit of the hearing aid. According to Korczak, Kurtzberg and Stapells (2005), when the compression circuit is active, a ceiling effect occurs for aided responses at higher intensities. Therefore, subtle variations in signal processing brought about by an increase or decrease in the number of channels, may not be accurately reflected if other processing schemes such as compression are active.

Correlation between speech identification scores and ACC measures

Non-parametric correlations were administered to achieve the forth objective which was to compare the SIS and ACC across different number of hearing aid channels. Spearman's correlation was done, considering the peaks in the ACC waveform namely; P2, N1' and P2' which was correlated with the speech identification score obtained using two, four and eight channel hearing aids. The mean scores obtained for the two, four and eight channels were 83, 84 and 81% with standard deviations of 8.48, 10.90 and 9.78 respectively. The difference between the unaided and the aided condition approached significance at $\chi^2(8) = 7.541$, p = 0.057. The same has been depicted in Figure 9.

A positive correlation between the amplitude of P2' with the four channel device and the speech identification scores ($\rho = 0.822$, $p < 0.05$). The failure of the two and eight channel devices to show similar findings could be attributed to the inherent redundancy of words compared to a syllable. Word discrimination improves with

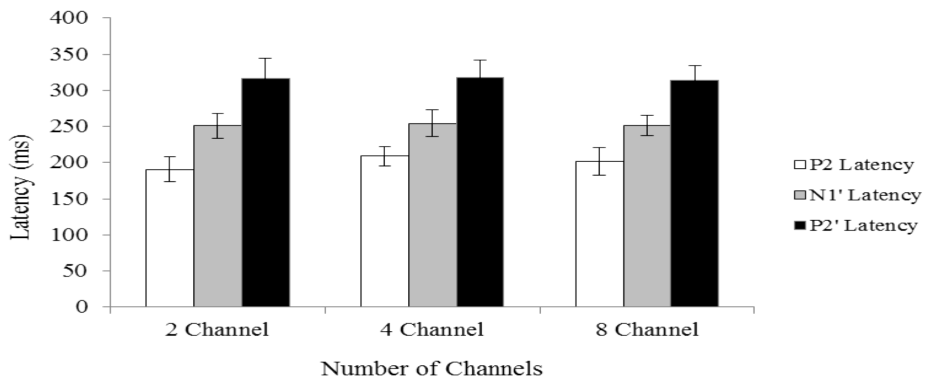


Figure 5: Mean latencies for the three aided conditions. Error bars denote one standard error around the mean.

increase in word length (Black, 1952). The speech identification material used in the present study consisted of monosyllabic words and the ACC stimulus was a single syllable /si/. The difference in redundancy between the two stimuli could have resulted in a lower correlation between the two. Subjective differences in perception through hearing aids with varying number of channels may also have be a reason. This needs to be clarified by further research in the same area.

In contrast to the present study, behavioral measures of discrimination and its correlation to CAEP have been assessed by Korczak et al., (2005). This study has found a correlation between the behavioral or discrimination score and the electrophysiological testing.

A reason for no significant changes seen between the number of channels and speech identification scores could be that an increase in the number of channels negatively affects the perception of diphthongs and vowels (Bor et al., 2008) and not to a large extent the perception of consonants according to Boothroyd et al., 1996. High Frequency Kannada Speech Identification Test (Mascarenhas, 2002) was used to assess the speech identification score.The high frequency content in the wordlist was contributed by consonants which were semivowels (/j/, /r/, /l/), stops (/t/, /p/, /k/), fricatives (s, /f/, /ʃ/) and the affricate /tʃ/. The consonants themselves could have sufficiently improved intelligibility without the need for depending on the vowels for perception. Therefore, only minor variations in the scores were seen across different number of channels.

Conclusion

When compared to individuals with hearing sensitivity within normal limits, those with cochlear hearing loss have delayed latencies on the ACC; amplitudes were found to be similar between the groups. For sloping cochlear hearing loss, with or without amplification the responses were similar. When the two, four and eight channel hearing aids were evaluated electrophysiologically through the ACC, the performance was equivalent. When the electrophysiological findings were correlated with the speech identification scores for each aid, ACC obtained with the four channel device was found to correlate with the SIS; but, there was no correlation found between the two and eight channel hearing aids.

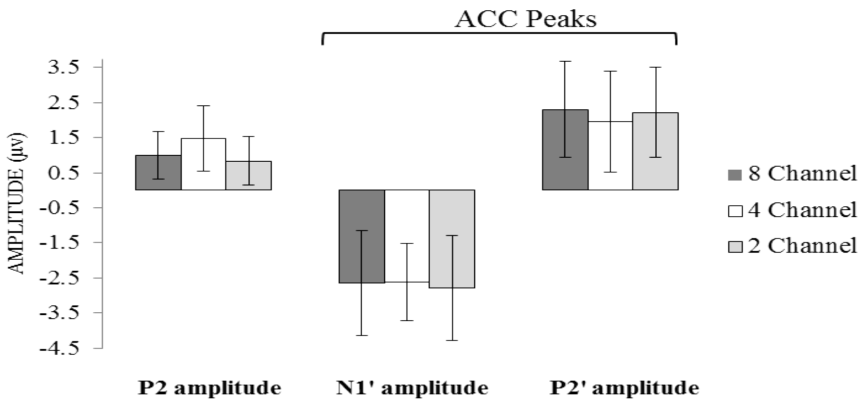


Figure 6: Mean amplitudes for the aided condition of two, four and eight channels. Error bars denote one standard error around the mean.

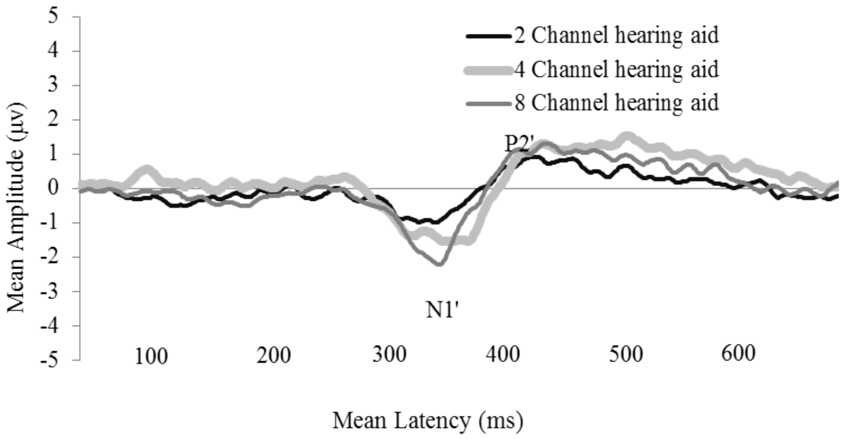


Figure 7: Grand mean waveform of the ACC obtained for the aided condition using the two, four and eight channel device.

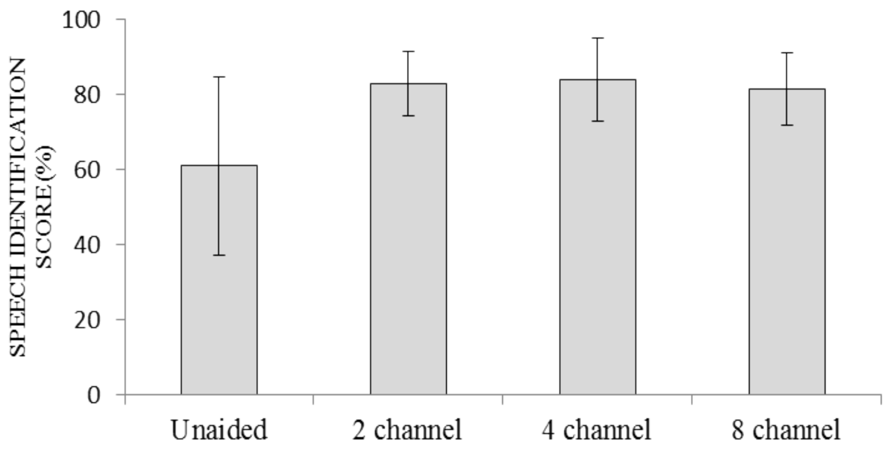


Figure 8: Mean speech identification scores obtained for different conditions.Error bars denote one standard error around the mean.

References

Agung, K., Purdy, S. C., McMahon, C. M., &Newall, P. (2006). The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *Journal of the American Academy of Audiology, 17*(8), 559-572.

Black, J. W.Â (1952). Accompaniments of word intelligibility. *Journal of Speech and Hearing Disorders, 17*, 409-418.

Boothroyd, A., Mulhearn, B., Gong, J., & Ostroff, J. (1996).Effects of spectral smearing on phoneme and word recognition. *Journal of Acoustical Society of America, 100,* 1807-1818.

Bor, S., Souza, P., & Wright, R. (2008). Multichannel compression: Effects of reduced spectral contrast on vowel identification. *Journal of Speech Language and Hearing Research, 51*, 1315-27.

Buus, S., & Florentine, M. (2001). Growth of loudness in listeners with cochlear hearing losses: Recruitment reconsidered. *Journal of the Association of Research in Otolaryngology, 3,* 120-139.

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

Demeester, K., van Wieringen, A., Hendrickx, J., Fransen, E., van Laer, L., Van Camp, G., & Van de Heyning, P. (2009). Audiometric shape and Presbycusis. *International Journal of Audiology, 48*, 222-2332.

Dillon, H. (2001). *Hearing aids*.Â Turramurra, Australia:Â Boomerang Press.

Easwar, E., Glista, D., Purcell, D. W., & Scollie, D. S. (2012). Hearing aid processing changes tone burst onset: Effect on cortical auditory evoked potentials in individuals with normal audiometric thresholds. *American Journal of Audi-*

ology, 21(1), 82-90.

Jyoti (2010).*Effect of number of channels of hearing aids on speech perception in different degrees of sloping hearing loss cases.* Unpublished Masters's Dissertation submitted to the University of Mysore, Mysore.

Korczak, P. A., Kurtzberg, D., & Stapells, D. R. (2005).Effects of sensorineural hearing loss and personal hearing aids on cortical event-related potential and behavioral measures of speech-sound processing. *Ear and Hearing, 26*(2), 165-185.

Margolis, R., & Saly, G. (2008). Distribution of hearing loss characteristics in a clinical population. *Ear and Hearing,29*(4), 524-532.

Martin, B. A., & Boothroyd, A. (2000). Cortical, auditory, evoked potentials in response to changes of spectrum and intensity. *Journal of the Acoustical Society of America, 107,* 2155-2161.

Martin, B. A., Tremblay, K. L., & Stapells, D. R. (2007).Principles and applications of cortical auditory evoked potentials. In R. F. Burkard, M. Don, & J. J. Eggermont (Eds.), *Auditory Evoked Potentials. Basic Principles and Clinical Application* (pp. 482-507). Philadelphia: Lippincott, Williams & Wilkins.

Mascarenhas, K. (2002). *A high frequency Kannada speech identification test (HF-KSIT).* Unpublished Masters's Dissertation submitted to the University of Mysore, Mysore.

Moore, B. C. J., (1998). *Cochlear hearing loss.* London:Whurr Publishers Ltd.

Oates, P. A., Kurtzberg, D., & Stapells, D. R. (2002).Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing.*Ear and Hearing, 23*, 399-415.

Polen, S. B. (1984). Auditory event related potentials. *Seminars in Hearing, 5,* 127-141.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function. *Journal of the Acoustical Society of America, 83*, 2322-2327.

Roeser R. J., & Clark, J. L. (2007).Pure tone tests. In R. J. Roeser, H. Hosford-Dunn, M. Valente (Eds.), *Audiology Diagnosis (2 Ed.)*(pp. 238-260). New York:Thieme Medical Publishers, Inc.

Rubina (2008).*Effects of Degree of Loss and Age on Speech Identification with Multichannel Hearing Aids*.Unpublished Masters's Dissertation submitted to the University of Mysore, Mysore.

Souza, P. E., &Â Boike, K. T. (2006).Combining temporal cues across channels: Effects of age and hearing loss.*Journal of Speech, Language and Hearing Research, 49,* 138-49.

Stelmachowicz, P. G., Kopun, J., Mace, A., Lewis, D. E., & Nittrouer, S. (1995). The perception of amplified speech by listeners with hearing loss: Acoustic correlates. *Journal of the Acoustical Society of America, 98*, 1388-1399.

Stevens, S. S. (1955). The measurement of loudness.*Journal of Acoustical Society of America, 27,* 815-829.

Tremblay K. L., Billings C. J., Friesen L. M., & Souza P. E. (2006). Neural representation of amplified speech sounds. *Ear and Hearing, 27*, 93-103.

Wall, L., Dalebout, S. D., Davidson, S. A., & Fox, R. A. (1991).Effect of hearing impairment on event-related potentials for tone and speech distinctions.*Folia Phoniatrica, 43*, 265-274.

Yetkin, F. Z., Roland, P. S., Christensen, W. F., & Purdy, P. D. (2004). Silent fMRI of tontopicity and stimulus intensity coding in human primary auditory cortex, *Laryngoscope, 114*, 512-518.

Yund, E. W., & Buckles, K. M. (1995). Multichannel compression hearing aids: Effect of number of channels on speech discrimination in noise. *Journal of the Acoustical Society of America, 97,* 1206-1223.

Zeng, F. G., & Turner, C. W. (1990).Recognition of voiceless fricatives by normal and hearing-impaired subjects.*Journal of Speech and Hearing Research, 33*, 440-449.

Development of Tone Burst ABR in Infants

[1]Jasmine Abdulrehman & [2]Mamatha N. M.

Abstract

Frequency-specific ABR is universally recommended to determine the thresholds in infants who do not pass newborn hearing screening. However important variable in the interpretation of ABR tracings, especially in premature infants, is the effect of maturation on the waveforms. The present study was carried out with the aim of studying the effect of age on tone burst evoked ABR in infants and to see the maturational changes of tone burst evoked ABR in infants. Infants were divided into two groups; group I with age range of 0 to 6 months (15 ears) and group II with age range of 6 to 12 months (15 ears). Wave V latency changes of 500 Hz and 4000 Hz at intensity levels of 70 dB nHL, 50 dB nHL, 40 dB nHL and 30 dB nHL were analysed in both the groups. There was significant correlations for wave V latency across age groups for 500 Hz at all the intensities, and for 4000 Hz correlation was present only for 70 dB nHL and 50 dB nHL. Significant difference in latency was observed within age groups across 2 frequencies at different intensity levels. Use of tone burst ABR is helpful in finding the frequency specific threshold in infants. Also by seeing the latency change, maturation of the auditory system can be easily understood. ***Keywords:*** *Latency, ABR, Maturation Change, Tone burst, Latency.*

Introduction

Auditory brainstem response (ABR) is an electrical potential and it is a complex response to particular types of external stimuli which represents neural activity generated at several anatomical sites. Electrical potentials recorded from the scalp in response to auditory signal were first described in 1939 by Davis and colleagues. Auditory evoked potentials can be divided into two categories: transient, or onset potentials and sustained potentials. Neural units generating transient potentials are onset sensitive, and respond to the onset of the stimulus. Jewett (1970) demonstrated that neural responses could be recorded from the brainstem pathways of cats. Jewett, Romano and Williston (1970) recorded ABR in humans and reported that responses consisted of 5 peaks occurring within 7 ms of stimulation.

ABR is an electrophysiological measure used for assessing hearing sensitivity in individuals for whom conventional behavioural method will be difficult to perform. It is also used for identification of neurological abnormalities of the eighth nerve and auditory pathways of the brainstem. Hearing screening of newborn infants began in 1960s (Downs & Sterritt, 1967). Clinical application of the ABR in children appeared over thirty years ago (Hecox & Galambos, 1974). New born auditory screening is a major clinical application of the ABR. The different stimuli used to record ABR are click, tone burst and speech. Click with duration of 100μs is most commonly used for ABR measurement. Click is a broad band stimulus which activates wide range of area in the basilar membrane. There is a lack of agreement among different investigators regarding the frequency region most important for generation of click ABR, i.e., whether it reflects activation of 1000 to 4000 Hz region, 4000 to 8000 Hz region or 2000 to 4000 Hz region (Balfour, Pillion & Gaskin, 1998; Eggermont & Don, 1980; Stapells, 1989). Click ABR corresponds best to the zone of hearing between 500 Hz to 4k Hz. Often ABR testing employs clicks to assess high frequency hearing and tone burst to assess low frequency hearing (Hood, 1998; Goldstein & Aldrich, 1999; Sininger & Cone-Wesson, 2002).

In persons with high frequency hearing loss, the portion of the cochlea contributing to click ABR is different from normal hearing individuals. It varies as a function of response component and stimulus intensity. Wave I reflects activity from basal region and wave V from apical region. Also at higher intensity levels, spread of activation is noticed towards the apex whereas at lower intensity levels, activation is towards the basal region. Tone burst ABR is preferred to elicit frequency specific response from the auditory system especially in infants and children. Tone burst signal have sufficiently rapid onsets to elicit ABR effectively and at the same time maintaining relatively well-defined frequency specificity. Several studies have compared pure-tone thresholds and tone burst-evoked ABR thresholds (Kodera, Yamane, Yamada & Suzuki, 1977; Munnerley , Greville, Purdy & Keith, 1991; Purdy & Abbas, 2002; Stapells, 2000; Suzuki & Horiuchi, 1977). The agreement between the two threshold measures suggests that, tone burst-evoked ABR thresholds can be used to predict the magnitude and configuration of hearing loss (Suzuki, Kodera & Kaga, 1982). Significantly lower signal-to-noise ratio Tone-burst ABRs causes narrower peak excitation on the basilar membrane and also results in a less synchronized neural response and an overall smaller number of neurons responding (Hall, 2006). It has been noticed that stimuli at different frequencies and levels can yield very different wave morphologies.

Hurley, Hurley and Berlin (2005) compared the 500 Hz

[1]Email: 30jasminejss@gmail.com,
[2]Lecturer in Audiology, Email: mamms_20@rediffmail.com

tone burst and click evoked ABR in 305 infants with conceptional age from 33 weeks to 74 weeks. Absolute latencies of wave V at 55, 35, and 25 dB nHL were measured. Data obtained from this study provided age equivalent norms for 500 Hz tone burst. They concluded that Wave V latency in response to 500 Hz tone burst decreased with age and did not stabilize by 70 weeks conceptional age.

To estimate the infants hearing sensitivity using behavioural response is a difficult task. The disadvantage of Behavioural Observation Audiometry is that, it is difficult to eliminate tester bias. Behavioural responses of infants and young children reaches extinction quickly without reinforcement, and a wide variety of responses are noted in youngsters.

Evoked potentials are essential portion of audiologic evaluation mainly in these difficult to test subjects. However, one of the commonly used traditional evoked potential is limited to click ABR only. Click stimuli provides an estimation of auditory functioning in a broad frequency range between 1000 to 4000 Hz. A complete evoked potential evaluation must contain reliable ear specific and frequency specific information. Frequency specific ABR using brief tone is a better choice to get the contour of hearing loss.

Frequency-specific ABR test is universally recommended to determine the thresholds in infants who do not pass newborn hearing screening (Joint Committee on Infant Hearing, 2000). However important variable in the interpretation of ABR tracings, especially in premature infants, is the effect of maturation on the waveforms. The effect of maturation in infants is significant, and with proper normative, misinterpretation of peak wave latencies can easily be avoided. Common paradigm to assess more frequency-specific hearing includes tone bursts with a Blackman windowing function (Gorga, Kaminski Jesteadt & Neely, 1989; Gorga, Kaminski, & Beauchaine, 1988). Several studies have been conducted to see the agreement between pure tone behavioural threshold and tone burst ABR threshold. These agreements suggest for evaluating the magnitude and configuration as it serves as a better choice than click ABR. Threshold information using tone burst can also be used for programming hearing aid.

Age-equivalent ABR norms are important when predicting hearing levels in neonates and infants. Delayed absolute latencies can often indicate other pathologies or unique conditions. Therefore, it is useful to have age equivalent norms for different frequency specific tone burst stimuli to reach a better interpretation. It would be helpful for fitting the hearing aid in children with hearing loss.

Therefore, the present study was carried out with the aim of studying the effect of age on tone burst ABR in infants; to see the maturational changes of tone burst ABR in infants by documenting wave V latency changes of 500 Hz and 4000 Hz TB; to compare difference in latency across 500 Hz and 4000 Hz at different intensity levels (70 dB nHL, 50 dB nHL, 40 dB nHL & 30 dB nHL). The study also aimed to compare the effect of intensity on wave V latency at each frequency separately.

Method

Participants

The study was conducted on 30 infants (30 ears) in the age range of 0 to 12 months. Thirty infants were further subdivided into 2 groups based on their age as 0 to 6 months (15 ears) and 6 to 12 months (15 ears). All the infants had normal otoscopic examination indicating absence of external and middle ear pathology. They were healthy with no symptoms of cold or ear discharge at the time of assessment. They had no complaints and prior histories of any high risk factors, neurological symptoms. All the infants had age appropriate minimum response levels in behavioural observation audiometry, normal outer hair cell functioning ensured by recording TEOAEs and normal hearing sensitivity ensured by recording ABR.

Test Environment

Testing was carried out in a sound treated room where ambient noise levels were within the specified limits as per ANSI S3.1 (1991). The test room was made comfortable enough for the infants in terms of temperature and light.

Procedure

Case history: Detailed information regarding the history of prenatal, natal and postnatal medical conditions was secured for all the infants. A detailed report regarding the auditory behaviour of the infants at home for various environmental sounds like calling bell, pressure cooker, whistle etc were obtained from the parents or caregivers.

High risk register: Medical reports were reviewed to make sure that all the infants were devoid of various risk factors and other medical conditions which can affect auditory system. The modified high risk register developed by Anitha and Yathiraj (2001) was administered to rule out the high risk factors for hearing impairment in infants.

Otoscopic examination: The visual examination of the ear canal and the tympanic membrane of the infant's ear were done using a hand held otoscope. This was done to rule out the presence of wax, foreign bodies in the ear canal and/ or tympanic membrane pathologies.

Behavioural observation audiometry (BOA): The behavioural responses (minimum response level) of the infants were observed in the free field condition using warble tones from 500 Hz to 4000 Hz separated in octaves and for speech stimuli. The lowest levels of presentation of each of the stimuli, at which the subject exhibited some sort of auditory behaviour was noted down.

Tympanogram and Acoustic reflex measurements: The tympanograms were recorded over a pressure range of +200 to - 400 daPa with probe tone frequency of 226 Hz or 1000 Hz. For infants till the age of 6 months, 1000 Hz probe tone was used and for infants above 6 months of age, 226 Hz probe tone frequency was used, with pump rate of 200 daPa /s. To see for the presence of acoustic reflex, Ipsi lateral acoustic reflex thresholds were obtained using broad band noise.

Transient evoked otoacoustic emissions (TEOAEs): TEOAEs were obtained using ILO - V6 instrument with a foam tip positioned in the external auditory canal so as to give a flat stimulus spectrum across the frequency range. Stimuli were clicks with a band pass filter encompassing 500 Hz - 6000 Hz. The duration of rectangular pulses (click) was 80μs. The responses considered were based on the reproducibility and signal to noise ratio (SNR). The overall SNR of greater than or equal to + 3 dB and reproducibility of greater than 50% were considered (Dijik & Wit, 1987) for the presence of otoacoustic emissions.

Auditory Brainstem Response: ABR was recorded using single channel for all the infants using IHS Smart EP system. Initially electrode sites were cleaned with the help of skin preparing gel. Electrodes were placed on the recording sites with the conducting paste and then fixing them with the help of surgical tape. It was ensured that the independent electrode impedance was roughly equivalent or < 5 k*Ohm* and inter electrode impedance was within 2 k*Ohm* at the start of the test. If click ABR wave V was clearly seen at 30 dB nHL or 40 dB nHL and if these results correlated with behavioural observation audiometry, then the infants were considered to have normal hearing sensitivity. ABR was recorded using 500 Hz and 4000 Hz tone bursts of 2-0-2 cycles with black man envelope at 4 intensity levels 70 dB nHL, 50 dB nHL, 40 dB nHL and 30 dB nHL for each subject in both the groups. Tone bursts were presented at the repetition rate of 11.1/s using alternating polarity through ER- 3A insert ear phones. At each intensity level, tone burst ABR was recorded twice for replicability.

Analysis

To arrive at the goal, quantitative and qualitative analysis was done. Qualitative analysis was done based on the simple visual inspection of wave V for 500 Hz and 4000 Hz tone burst at 70 dB nHL, 50 dB nHL, 40 dB nHL and 30 dB nHL for each subject in both the groups. Quantitative analysis was done by obtaining wave V latency values for 500 Hz and 4000 Hz tone burst Sfrequencies at 70 dB nHL, 50 dB nHL, 40 dB nHL and 30 dB nHL for both the age groups. The obtained data for two TB frequencies and 4 intensities across age groups were compared using appropriate statistical analysis.

Results and Discussion

Qualitative Analysis of Tone Burst ABR Wave forms

Tone burst ABR waveforms which are recorded are qualitatively analysed separately for each tone burst frequency and for each age group. The ear from which tone burst ABR recorded was randomly selected.

Qualitative Analysis of 500 Hz tone burst ABR waveforms in Group I: Qualitative analysis of 500 Hz tone burst evoked ABR obtained in group I (15 infants- 15 ears) at different intensity levels revealed that: Wave V could be recorded in all 15 infants (100%) upto 40dB nHL but it could be recorded till 30 dBnHL only in 10 (66.6%) infants and the morphology of the wave recorded using 500 Hz tone burst was poorer with decreasing intensity.The 500 Hz tone burst ABR wave form at different intensity levels recorded in one of the infants from group I is given in Figure 1

500 Hz tone burst at different intensity levels in one of the infants of Group I: Qualitative analysis of 4000 Hz tone burst ABR waveforms in Group I: Qualitative analysis of 4000 HZ tone burst ABR obtained for group I at different intensity levels revealed that Wave V could be recorded in 15 infants (100%) up to 40 dB nHL and could be recorded only in 12 infants (80%) till 30 dB nHL. The 4000 Hz tone burst ABR wave recorded at different intensity levels in one of the infants from group I is given in Figure 2

It can be observed from the above figures (1 & 2), that the morphology of the tone burst ABR becomes better and clearer with increasing frequency of the stimulus. Morphology of the 4000 Hz tone burst evoked ABR waveforms was better compared to 500 Hz tone burst evoked ABR and the waveforms of 500 Hz tone burst ABR had relatively broader peak.

Qualitative analysis of 500 Hz tone burst ABR wave forms in Group II: Analysis of 500 Hz tone burst ABR obtained from Group II (15 infants- 15 ears) at different intensity levels revealed that wave V could be recorded in all 15 infants (100%) upto 40 dB nHL but only in 11 infants (73.3%) till 30 dB nHL. The 500 Hz tone burst ABR wave at different intensity levels in one of the infants from Group II is given in Figure 3

Qualitative analysis of 4000 Hz tone burst ABR wave

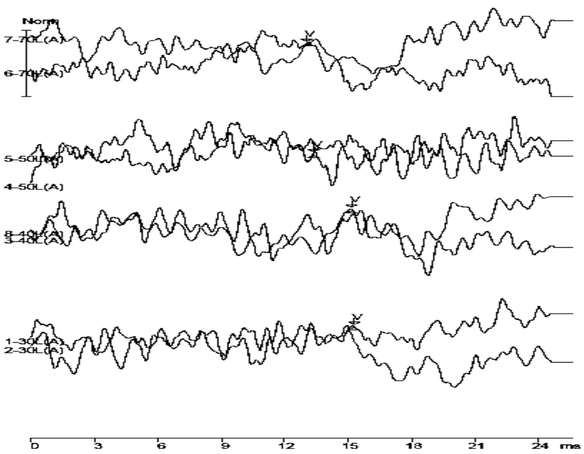


Figure 1: Waveforms showing wave V latency recorded for 500 Hz tone burst at different intensity levels in one of the infants of Group I.

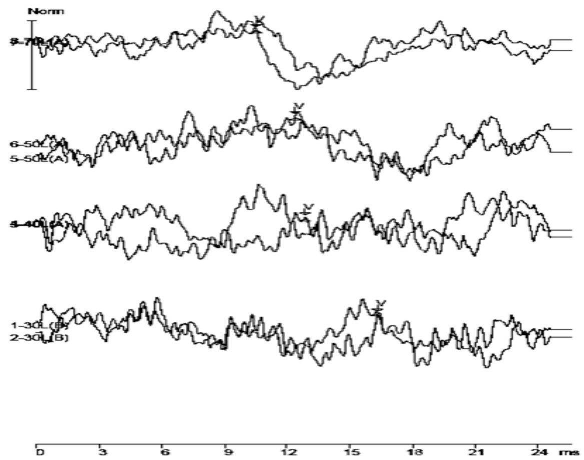


Figure 3: Waveform showing wave V latency recorded for 500 Hz tone burst at different intensity levels in one of the infants from Group II.

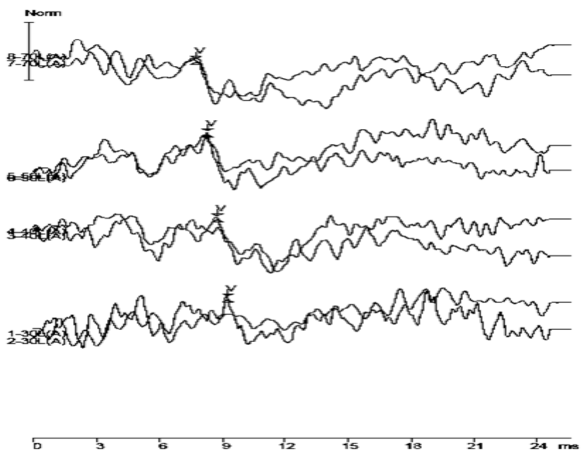


Figure 2: Waveforms showing wave V latency recorded for 4000 Hz tone burst at different intensity levels in one of the infants from Group I.

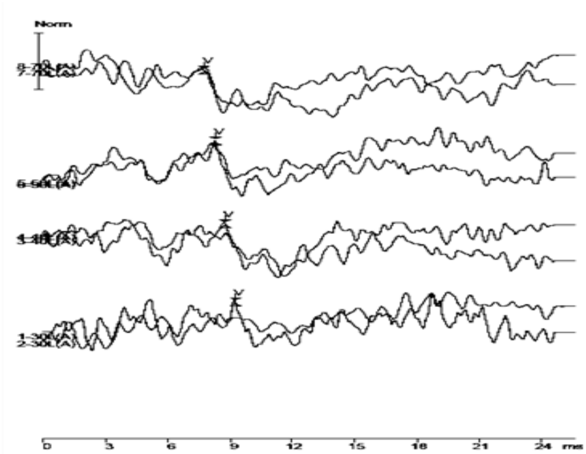


Figure 4: Waveform showing wave V latency recorded for 4000 Hz tone burst at different intensity levels in one of the infants from Group II.

forms in Group II: Qualitative analysis of 4000 Hz tone burst ABR obtained from Group II (15 infants- 15 ears) at different intensity levels revealed that Wave V could be recorded in 15 infants (100%) upto 40 dB nHL and 14 infants (93.3%) had wave V till 30 dB nHL. The 4000 Hz tone burst ABR wave form recorded at different intensity levels in one of the infants from Group II is given in Figure 4.

As observed in Group I, the morphology was poorer with broader peak at 500 Hz tone burst ABR when compared to 4000 Hz tone burst ABR. It can also be noted that wave morphology was better in group II than in group I.

Quantitative Measurements

Quantitative measurements were done to compare the difference in wave V latency with change in age, frequency and intensity. Tone burst ABR obtained from 15 infants in group I and 15 infants in group II were compared in terms of the latency of the wave V. Wave V latency of TB ABR recorded using 500 Hz and 4000 Hz at 70 dB nHL, 50 dB nHL, 40 dB nHL and 30 dB nHL were analysed using Statistical Package for the Social Sciences (SPSS) version 16 software.

Comparison of wave V latency values at different intensities, across age groups using different tone burst frequencies: The mean, standard deviation and range for wave V latency at each age group independently at different intensities using different tone burst frequencies.

As it can be seen in Table 1, there is a specific trend seen in mean wave V latency values. It can be observed that, the mean latency value is higher for group I (0-6 months) compared to group II (6- 12 months). It can also be noted from Table 1 that, the mean latency is

Table 1: Mean, Standard deviation (S.D) and Range of Wave V latency in 'ms' across age groups for different tone burst frequency

Toneburst Frequencies	Intensities (dB nHL)	Wave V latency (ms) Group I Mean	S.D	Range	Group II Mean	S.D	Range
500 Hz	70	12.20	0.85	11.05-13.15	10.28	0.63	8.97-11.10
	50	13.58	0.67	12.60-15.45	11.40	0.58	10.53-12.46
	40	14.59	0.83	13.50-15.70	12.30	0.58	11.50-13.35
	30	16.00	0.97	14.75-17.50	13.39	0.91	12.55-15.60
4000 Hz	70	7.52	0.42	6.90-8.35	6.92	0.38	6.50-7.70
	50	8.26	0.49	7.20-9.30	7.50	0.39	6.85-8.20
	40	8.77	0.68	7.75-10.25	8.20	0.58	7.25-9.20
	30	9.72	1.53	8.30-12.80	8.76	0.42	8.12-12.80

higher at low frequency, i.e. 500 Hz and lower at 4000 Hz in both the groups. As the intensity level decreases from 70 dB nHL to 30 dB nHL, the mean latency value was found to increase in both the groups.

Mixed analysis of variance (Mixed ANOVA) was done to see the significant interaction across age groups, tone burst frequencies and intensities for wave V latency measures. The results revealed a significant interaction across age groups [$F(1, 17) = 52.97$, $p < 0.05$], across frequencies [$F(1, 17) = 1010.15$, $p = 0.001$], across intensities [$F(3, 51) = 147.45$, $p < 0.05$], across age groups and frequencies [$F(1,17) = 26.04$, $p < 0.05$], across age groups and intensities [$F(3,51) = 1.341$, $p < 0.05$] and across frequencies and intensities [$F(3, 51) = 12.227$, $p < 0.05$]. There is no significant interaction across age groups, frequencies and intensities [$F(3, 51) = 0.301$, $p > 0.05$].

Effect of intensity on wave V latency at each tone burst frequency on each age group: As mixed ANOVA showed significant interaction of wave V latency measures across intensities, repeated measure ANOVA was done for each age group independently to see the effect of intensity on wave V latency at each tone burst frequency. There was significant difference in wave V latency across intensities for Group I at both 500 Hz [$F(3, 27) = 54.03$, $p < 0.05$] and 4000 Hz [$F(3,33) = 29.27$, $p < 0.05$] tone burst frequencies. Group II also showed a significant difference in wave V latency across intensities for 500 Hz [$F(3,39) = 140.76$, $p < 0.05$] and 4000 Hz [F (3,30= 50.57, $p < 0.05$)] tone burst frequencies.

As repeated measure ANOVA showed significant difference in wave V latency across intensities when 500 Hz and 4000 Hz tone burst frequency was used for group I and group II, further analysis was done using Bonferroni's multiple pair wise comparison test, to see between which two intensities, latency of wave V differ significantly. The results of the Bonferroni's multiple pair wise comparison test show that when 500 Hz tone burst frequency was used there was intensity effect seen with significant difference across all the 4 different intensities in both group I and II ($p<0.05$). Similarly when 4000 Hz tone burst frequency was used, there was significant difference across all the 4 different intensity levels from 70 dB nHL to 30 dB nHL in both age group ($p<0.05$).

In the present study there was effect of intensity on wave V latency at 500 Hz and 4000 Hz on Group I and Group II. It was found that there was an increase in latency with decreasing intensity from 70 dB nHL to 30 dB nHL. Similar trend is observed in both Group I and Group II at both frequencies (500 Hz & 4000 Hz). Study done by Werff, Prieve and Georgantas (2009) compared latency intensity function in infants and adults. They have reported that the slopes of the adult latency intensity functions at 2000 Hz (61 μs/dB) and 4000 Hz (45μs/dB) were similar or identical to those for infants at the same frequencies. They also observed that at 500 Hz, the adult slope was slightly shallower (80 μs/dB) than the infant slope.

Beattie (1998) studied ABR on 40 normal hearing subjects for both air-conducted and bone-conducted clicks at intensities of 55, 40, 30, 20, and 10 dB SL. Both bone conduction and air conduction click exhibited increase in latency as the intensity decreased to 10 dB SL. Similarly Fausti, Olson, Frey, Henry and Schaffer (1993) studied latency intensity function using 8 kHz, 10 kHz and 14 kHz tone burst ABR in 14 adults. Result showed shift in response latency for all the frequencies, shift of 0.02 ms to 0.03 ms/ dB was observed.

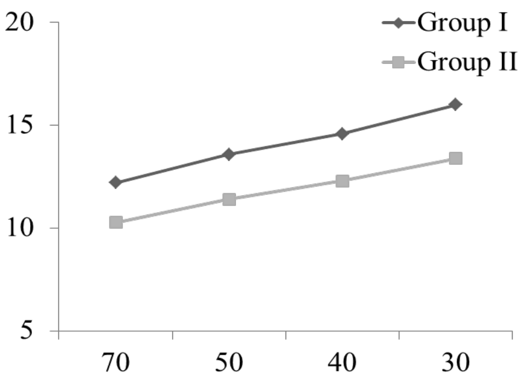


Figure 5: Changes in wave V latency at 500 Hz tone burst in both Group I and Group II.

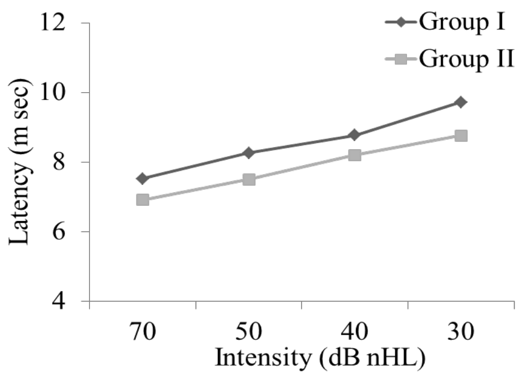


Figure 6: Changes in wave V latency at 4000 Hz tone burst in both group I and group II.

Effect of tone burst frequency on wave V latency at each age group at each intensity: As mixed ANOVA showed significant interaction of wave V latency measures across tone burst frequencies, repeated measure ANOVA was done to see the effect of tone burst frequency on wave V latency at each age group. Results showed that, there was significant difference in wave V latency between tone burst frequency across intensities for Group I [F (1, 8) = 1095, $p < 0.05$] and Group II [F (1, 7) = 275.34, $p < 0.05$].

Further analysis was done using paired t test, to see at which intensities, latency of wave V differ significantly between the two tone burst stimuli. The results of the paired t test showed that there was a significant difference between wave V latency when 500 Hz and 4000 Hz tone burst were used at all the 4 different intensities from 70 dBnHL to 30 dBnHL for both Group I and Group II ($p<0.05$). It was found that wave V latency was significantly more for 500 Hz tone burst compared to 4000 Hz tone burst in both, Group I and Group II. Studies related to latency changes with frequency have reported the similar results. Suzuki, and Horiuchi (1977) reported that when toneburst stimuli is presented in quiet, the latency and amplitude changes are more for low frequency tones.

Similarly, Gorga et al. (1988) evaluated tone burst ABR in 20 normally hearing subjects for a wide range of frequencies (250-8000 Hz) and levels. Findings suggest that, there was decrease in wave V latencies with increase in both frequency (250 - 8000 Hz) and levels (20 to 100 dB SPL). Reason for increase in wave V latency with decreasing frequency was given by Gorga et al. (1988), reported that decrease in wave V latency with increase in frequency while level is held constant could be due to difference in stimulus rise time. States that increase in latency might be due to longer rise time used in low frequency, and also as frequency decreases, place of maximum excitation shifts towards the apical end of the cochlea. Slop of latency frequency function decreases as level is increased, indicates that latencies are less dissimilar for different frequencies at high levels. They concluded that this pattern could be due to greater effective spread of excitation for low frequency stimuli than high frequency stimuli. Likewise decrease in wave V latency with increase in level while frequency is held constant could result from a basal spread of excitation.

Effect of age on wave V latency at each tone burst frequency: As there was significant interaction seen between age groups when mixed ANOVA was done, to see the effect of age on wave V latency at different tone burst frequencies and at each intensity levels, MANOVA was done for both age groups. Results showed that there was significant difference in wave V latency between groups when 500 Hz [F (1, 17) = 57.20, $p < 0.05$] and 4000 Hz [F (1,17)= 13.80, $p < 0.05$] tone burst was used. Further analysis of MANOVA using pair wise comparison was done to see at which two intensities, latency of wave V differ significantly across groups for both tone burst frequencies. The results showed that when 500 Hz tone burst was used there was a significant age effect seen on latency at all intensities ($p<0.05$). When 4000 Hz tone burst was used, there was age effect seen with significant difference between Group II and II at 70dB and 50dB intensity levels but there was no significant age effect seen at intensity levels of 40dB and 30dB ($p < 0.05$).

Result of the present study shows that with increase in age there was change in 500 Hz wave V latency in all the intensity levels (70 dB nHL, 50 dB nHL, 40 dB nHL & 30 dB nHL), where as in 4000 Hz change in wave V latency was observed only for higher intensities (70 dB nHL and 50 dB nHL), not for lower intensities (40 dB nHL & 30 dB nHL) or when it reaches near threshold changes in latency with age decreases. Also it can be observed that decrease in wave V latency was more for 500 Hz TB with increasing age compared to 4000 Hz as shown in the Figure 5 and 6.

The results of the present study support the study done by Teas, Klein and Kramer (1982) using filtered click stimuli of 1 kHz, 2 kHz, 4 kHz and 8 kHz recorded ABR in infants from 4 to 60 weeks. They reported wave v latencies at 60 weeks of gestational age age match adult values only for 1 kHz at 50 dB and not at 30 dB, largest latency difference between adult and infants were observed for higher frequencies (2, 4 & 8 kHz) at 60 weeks of age.

Similarly another study done by Ponton, Eggermont, Coupland and Winkelaar (1991) reported that at low frequency TB wave V latency stabilizes by 23 to 24 months post birth. Werff et al. (2009) reported that, infants at 3 months of age showed longer latency difference compared to adults at 500 Hz tone burst stimuli. Study done by Hurley et al. (2005) in infants support the present result where they showed that, 500 Hz tone burst latency decreases with age and not stabilized by 70 weeks of conceptional age. The study also showed that, wave V latency for the 500 Hz AC-TB ABR followed a predicted decrease in latency with age across the age range from 33 to 74 weeks of conceptional age mainly at 55 and 35 dB nHL.

Werner, Folsom and Mancl (1993) also showed using tone burst of 1, 4, and 8 kHz, larger decrease in latency was observed between 6 months and adults and smaller decrease between 3 months to 6 months. Similar to the present results at 4 kHz, there was a significant age and level interaction, also significant difference between 3 months and older listeners at 20 dB HL and significant difference between infants and adults at 30 and 40 dBHL was noticed. Reason for less latency shift with increasing age in the present study could be similar to the study done by Ponton et al. (1991) studied frequency specific maturation of the eight nerve and brainstem auditory pathway using derived ABR in infants and young children (2 weeks post term to 9 years) and adults. Both full term and premature infants ranging from 29 to 42 weeks were included for the study. They reported that mid frequency channels of the VIII nerve and auditory brainstem matured faster and earlier than responses from either very low or very high frequency channels.

Conclusions

Finally it can be inferred from the present study that changes in age causes shift in wave V latencies for 500 Hz tone burst at all the intensities and for 4000 Hz shift was observed only for 70 dB nHL and 50 dB nHL. Compared to 500 Hz tone burst, 4000 Hz tone burst gave better morphology and lower wave V latency. This study supports that use of tone burst ABR is helpful in finding frequency specific threshold in infants. Also by seeing the latency change maturation of the auditory system can easily be understood. Present study was conducted in 30 infants (30 ears) only. The study can be conducted further on more number of infants. Comparison was made only on 2 age groups, can be conducted considering higher age groups and also tone burst frequency taken for the present study is limited to 500 Hz and 4000 Hz, further study using more frequencies can be carried out.

References

Anitha. T. (2001). *Modified high risk register for professional and non-professional formulation and its efficacy.* Un published independent project. University of Mysore, Mysore.

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

Balfour, P. B., Pillion, J. P., & Gaskin, A. E. (1998). Distortion product otoacoustic emission and auditory brain stem response measures of pediatric sensorineural hearing loss with islands of normal sensitivity. *Ear and Hearing, 19,* 463-472.

Beattie, R. C. (1998). Normative wave V latency intensity functions using the ear tone 3A Insert earphone and the Radioear B71 bone vibrator. *Scandinavian Audiology, 27*, 120-126.

Dijik, P., & Wit, H. P (1987). The occurrence of click evoked otoaoustic emission in normal hearing ears. *Scandinavian Audiology, 16*, 62-64.

Downs, M. P., & Sterritt, G. M. (1967). A guide to newborn and infants hearing screening programs. *Archives otolaryngology, 85*, 37-44.

Eggermont, J. J., & Don, M. (1980). Analysis of the lick evoked brainstem potentials in human using high pass noise masking. II. Effect of click intensity. *Journal of the Acoustical Society of America, 68*, 1671-1675.

Fausti, S. A., Olson, D. J., Frey, R. H., Henry, J. A., & Schaffer, H. I. (1993). High frequeny tone burst evoked ABR latency intensity functions. *Scandanavian Audiology, 22*(1), 25 -33.

Goldstein, R. & Aldrich, W. M. (1999). *Evoked Potential Audiometry*. Boston: Allyn and Bacon.

Gorga, M. P., Kaminski, J. R., Beauchaine, K. L., Jesteadt, W., & Neely, S. T. (1989). Auditory brainstem responses from children three months to three years of age: Normal patterns of response II. *Journal of Speech and Hearing Research, 32,* 281-288.

Gorga, M. P., Kaminski, J. R., & Beauchaine, K. L. (1988). Auditory brainstem responses from children three months to three years of age: normal patterns of response. II. *Journal of Speech Hearing Research, 32,* 281-288.

Hall, J. W. (2006). *New handbook of auditory evoked responses*. Boston, Allyn and Bacon.

Hecox, K. & Galambos, R. (1974). Brainstem auditory evoked response in human infants and adults.

Archive Otolaryngology, 99, 30-33.

Hood, L. J. (1998). *Clinical application of the auditory brainstem response*. San Diego: Singular publishing group.

Hurley, R. M., Hurley, A., & Berlin, C. I. (2005). Development of low frequency tone burst versus the click auditory brainstem response. *Journal of the American Academy of Audiology, 16*, 114-121.

Jewett, D. L., Romano, M. N., & Williston, J. S. (1970). Human auditory evoked potentials: possible brain stem components detected on the scalp.*Science, 167*, 1517-1518.

Jewett, D. L. (1970) Volume-conducted potentials in response to auditory stimuli as detected by averaging in the cat. *Electroencephalography and Clinical Neurophysiology, 28*, 609-618.

Joint Committee on Infant Hearing (2007). Year 2007 position statement and guidelines for early hearing detection and intervention programs. *Journal of Paediatrics, 120*, 898.

Kodera, K., Yamane, H., Yamada, O., & Suzuki, J. I. (1977) Brain stem response audiometry at speech frequencies. *Audiology, 16,* 469-479.

Munnerley, G. M., Greville K. A., Purdy, S. C., & Keith, W. J. (1991). Frequency-specific auditory brainstem responses relationship to behavioural thresholds in cochlear-impaired adults. *Audiology, 30*, 25-32.

Ponton, C. W., Eggermont, J. J., Coupland, S. G., & Winkelaar (1991). Frequency specific maturation of the eighth nerve and brain stem auditory pathway: evidence from derived auditory auditory brain stem responses (ABR's). *Journal of the Acoustical Society of America, 91*(3), 1576-1586.

Purdy, S. C., & Abbas, P. J. (2002). ABR threshold to tone burst gated withBlackman and linear windows in adults with high- frequency sensorineural hearing loss. *Ear and hearing, 23*(4), 358-368.

Sininger. Y.S., & Cone- Wesson, B. (2002). Identification of neonatal hearing impairment: Infants with hearing loss. *Ear and Hearing, 1*(5), 488-507.

Stapells, D. R. 1989. Auditory brainstem response assessment of infants and children. *Seminars in Hearing, 10*, 229-251.

Stapells, D. R., (2000). Threshold Estimation by the Tone Evoked Auditory Brainstem Response: A Literature Meta Analysis. *Journal of Speech Language pathology and Audiology, 24*(2), 74-83.

Suzuki, T. & Horiuchi, K. (1977). Effect of high-pass filter on auditory brain stem responses to tone pips. *Scandinavian Audiology, 6*, 123- 126.

Suzuki, T. & Horiuchi, K. (1977). Effect of high-pass filter on auditory brainstem responses to tone pips. *Scandinavian Audiology, 6*, 123-126

Suzuki, J., Kodera, K., & Kaga, K. (1982). *Auditory evoked brainstem response assessment in otolaryngology. Annals of the New York Academy of science, 388,* 487-513.

Teas D. C, Klein A. J, & Kramer S. J. (1982). An analysis of auditory brainstem responses in infants. *Hearing Research, 7*, 19-54.

Werff, K. R., Prieve, B. A. &. Georgantas, L. M. (2009). Infant air and bone conduction tone burst auditory brainstem responses for classification of hearing loss and relationship to behavioral threshold (2009). *Ear and Hearing, 30*, 350-368.

Werner, L. A., Folsom, R. C., & Mancl, L. R. (1993). The relationship between auditory brainstem response and behavioural threshold in normal hearing. *Hearing Research, 68*, 131-141.

Aided Acoustic Change Complex in Individuals with Sensorineural Hearing Loss

[1]Jobish T. J. & [2]Sreeraj K.

Abstract

Multiple overlapping P1-N1-P2 responses can be elicited in response to naturally produced speech stimulus. This overlapping neural response represents the acoustical changes within speech stimulus. These potentials are known as Acoustic Change Complex (ACC). Research evidences clearly indicates that ACC can be reliably recorded in normal hearing individuals. But, there is lack of evidences in individual with sloping sensorineural hearing loss with different duration and degree of hearing loss and its effect on unaided and aided ACC. Hence, the aim of the present study was to investigate the effect of degree and duration of sloping sensorineural hearing loss and digitally amplified speech by hearing aid on ACC. The result of the study indicated that a clear ACC can be successfully recorded in aided condition compared to unaided condition. The present study also showed that as the degree and duration of hearing loss increases, there is a change in unaided and aided ACC compared to lesser degree and duration of hearing loss.

Keywords: *Acoustic Change Complex (ACC), Sloping sensorineural hearing loss, Unaided ACC, Aided ACC.*

Introduction

Sensorineural hearing loss (SNHL) is probably the most common form of hearing loss and these type of hearing loss not only lead to elevation of threshold for detection of sound, but, also the affects the way in which sound is perceived. The perception of speech in individuals with SNHL is also dependent on the configuration of hearing loss (Coughlin, Kewely-Port & Humes, 1998; Dubno, Dirks & Schafer, 1987; Sher & Owens, 1974). In SNHL, different audiometric patterns like flat, raising, sloping etc. are evident (Pittman & Stelmachowicz, 2003). One of the most common management options for this type of permanent hearing loss are hearing aids. Hearing aids can compensate sensorineural hearing loss by amplifying sound. Despite adequate amplification of sound by hearing aid, the person with cochlear hearing loss continues to report unclear and distorted speech. The effectiveness of hearing aid in persons with SNHL also depends on the ability of the central auditory system to represent and integrate the spectral and temporal information delivered by the hearing aid, other than hearing aid related factors.

The individuals with SNHL avoid the use of hearing aids or adjust to the listening situations without the use of hearing aids. This kind of avoidance and adjustments for a longer period of time can lead to auditory deprivation (Silman, Gelfand & Silverman, 1984). Hence, there is interest in examining the neural representation (by hearing aid) of speech cues and amplified speech cues in people with hearing loss. It will be still more interesting, if it is possible to get knowledge about the different durations of hearing loss and its effect on neural representation of speech cues and amplified speech cues (by hearing aid) in people with hearing loss.

[1]Email: tjjobish88@gmail.com,
[2]Lecturer in Audiology, Email: sreerajkonadath@aiishmysore.in

The P1-N1-P2 complex was the first tool for assessing the neural representation of sound in population with and without hearing loss. The P1-N1-P2 complex is an auditory evoked potential that is characterized by a positive peak (P1), followed by a negative peak with a latency of about 100 milliseconds after stimulus onset (N1), followed by a positive peak called P2. These peaks reflect neural activity generated by multiple sources in the thalamic-cortical segment of the central auditory system. Multiple overlapping P1-N1-P2 responses can be seen in response to naturally produced speech stimuli like syllables, words etc. If we consider multiple overlapping responses to a syllable, the first P1-N1-P2 complex reflects the onset of the consonant whereas the second P1-N1-P2 response reflects the consonant vowel (CV) transition. These complex waveform patterns were shown to reflect acoustic changes, from silence to sound (onset of consonant) and the CV transition (from consonant to vowel) (Ostroff, Martin & Boothroyd, 1998). Martin and Boothroydin (1999) termed this cortical evoked response as 'Acoustic Change Complex' (ACC). ACC and other cortical potentials like mismatch negativity (MMN), P300 give information on the auditory discrimination ability of individual. But, to elicit MMN and P300, individual's cooperation is required whereas for ACC, the same is not mandatory. Moreover, the average amplitude of ACC is 2.5 times larger compared to MMN in normal hearing individuals (Martin & Boothroyd, 1999). These advantages of ACC over other potentials in normal hearing individuals gives more interest to probe into study of effect of hearing impairment and amplification on cortical representation of acoustic changes within speech sounds. It is still more imperative to study effect of sloping sensorineural hearing loss (SNHL) on ACC, reason behind is that individual with sloping SNHL has more problem with perception of speech because the higher frequency speech sounds contributes primarily to the speech intelligibility. Hence, the high frequency speech sounds like fricatives within a syllable can be used to elicit ACC. Very few researchers gave more interest to probe more into it (Tremblay, Kalstein, Billings & Souza, 2006, Tremblay, Friesen, Martin & Wright, 2003).

Tremblay et al. (2006) found that /shee/ and /see/ each elicit distinct ACC responses. The first negative peak, signaling the onset of the consonant, was not significantly different for the /shee/ and /see/ stimuli. However, the second P1-N1-P2 complex (N345 and P413), presumably reflecting the CV transition, occurred significantly earlier when evoked by the /shee/ stimulus than when evoked by /see/. Onset of the vowel in /shee/ was 30 milliseconds earlier than the onset of the vowel portion in /see/. This 30-millisecond difference appears to correspond to the 30-millisecond latency difference between the negative peak N345 elicited by the /shee/ stimulus and the negative peak N375 elicited by the /see/ stimulus.

Tremblay et al. (2003), examined ACC patterns in normal-hearing young adults, evoked by two different speech stimuli, /shee/ and /si/. These particular stimuli were chosen because they share similar acoustic features and are frequently confused by listeners with hearing loss. /shee/ and /si/ are similar in that they are fricatives and different in that (1) /shee/ and /si/ differ by place of articulation, (2) the fricative portion of /shee/ contains lower spectral energy than the fricative portion of /si/, and (3) the fricative portion of /shee/ is shorter in duration than the fricative portion of /si/. Karthik and Vanaja (2005) demonstrated that ACC is an electrophysiological index of speech discrimination in adult and children. The study showed that ACC could be recorded in all adult subjects and there was a significant difference between N1P2 amplitudes between two stimuli indicative presence of ACC response.

All these research evidences suggest that ACC can be recorded successfully in adult normal hearing individuals. So, there is a need to know how successfully the aided ACC can be recorded in person with cochlear hearing loss and to correlate the findings to various degree and duration of hearing loss. In a study by Tremblay, Kalstein, Billings and Souza (2006) recorded ACC in adult hearing aid users with mild to severe degree of sloping sensorineural hearing loss for two consonant-vowel (CV) syllables (/shee/ and /see/). The result of the study showed that /shee/ and /see/ elicited different waveforms in terms of latency and amplitude. This finding indicated different neural detection of CV transitions (indicated by the presence of a P1-N1-P2 response) for /shee/ and /see/.The latency of the second N1 in the evoked cortical neural response coincided in time with the onset of the vowel in the /shee/ and /see/ syllable. This finding is in accordance with the study done by Tremblay, Billings, Friesen, and Souza in 2006. To critically evaluate, the study demonstrated the coding of acoustic changes in the subjects with mild to severe sloping sensorineural hearing loss in aided condition but did not studied the effect of mild to severe degree of SNHL on ACC. Moreover, duration of sloping SNHL along with the degree of hearing loss could also have impact on unaided and aided ACC.

Interestingly, the further studies in the literature also failed to make an attempt to study effect of degree and duration of sloping SNHL on unaided and aided ACC. This leads to the initiation of the current study.

Firstly**,** the persons with sloping SNHL have problems in perceiving acoustic changes within speech sounds (mainly the syllables having high frequency consonant and low frequency vowel combination). The ability to perceive these acoustic changes is very important for normal perception of speech. How well the acoustic changes within the speech stimuli are coded in the cortical level can be studied electro physiologically using ACC. So, there is a need to study the ACC in individual with varying degree of sloping hearing loss to see cortical representation of acoustic changes within speech sounds.

Secondly, the amount of benefit the hearing aid provides varies greatly among individual with sloping sensorineural hearing loss. The factors contributing could be that the hearing aid is incapable to process spectral and temporal information properly or could be the inability of the peripheral and central auditory system in processing acoustic changes within speech sounds provided through hearing aid. Hence, there is also a need to study the neural representation of (cortical representation) acoustic changes within speech sounds through a hearing aid in individual with sloping SNHL across the various durations of losses.

Thirdly, the degree and duration of hearing loss could be the factors that affect the cortical representation of acoustical changes within speech sounds in individual with sloping SNHL. Hence, there is a need to compare the effect of different degrees across various duration of hearing loss with and without the use of amplification, and study the difference in cortical representation of speech. Therefore the aim of the study was to investigate the effect of degree and duration of sloping sensorineural hearing loss and digitally amplified speech on Acoustic Change Complex.

Method

Participants

Total of 19 individuals with unilateral or bilateral sloping sensorineural hearing loss who are the naive hearing aid users were selected as participants of the study. The participants were in the age range of 19 to 55 years

(17 males and two females with a mean age range of 41.31 years) and the hearing loss was post lingual in onset. A total of 21 ears with sloping sensorineural hearing loss were selected for the recording of aided and the unaided ACC who has less than or greater than two years of duration hearing loss. The slope of the audiogram was defined based on the occurrence of the thresholds at equal or successively higher levels from 250 to 8000 Hz and the difference between thresholds at 250 and 8000 Hz was always >20 dB (Pittman & Stelmachowicz, 2003). The degree of the slope of audiogram was calculated based on PTA_1 (average of the pure tone thresholds at 500Hz, 1 kHz, 2 kHz) and PTA_2 (average of the pure tone thresholds at 1kHz, 2 kHz, 4 kHz). All 27 ears were divided into two groups, group one had 10 ears with minimal to moderate sloping cochlear hearing loss and group two had 11 ears with moderate to severe sloping sensorineural hearing loss. The number of ears in the group I and group II were categorized into four subgroups based on duration of hearing loss; minimal to moderate sloping sensorineural hearing loss (SNHL) with duration of hearing loss less than two years (subgroup A), minimal to moderate sloping SNHL with duration of hearing loss greater than two years (subgroup B), moderate to severe sloping SNHL with duration of hearing loss less than two years (subgroup C), moderate to severe sloping SNHL with duration of hearing loss greater than two years (subgroup D) as shown in the Table 1.

Table 1: Number of ears selected in each subgroup based on duration of hearing loss

Degree of Sloping cochlear hearing loss	Duration of hearing loss(Number of ears)	
	<2 years	*>2 years*
Minimal to Moderate	6	4
Moderate to Severe	4	7

Test Procedure

Procedure involved four phases: 1) Selection of participants, 2) Selection of hearing aid, 3) Hearing aid fitting, 4) Acquiring aided and unaided Acoustic change complex

Phase 1: Selection of Participants

The following audiological tests were carried out for participant selection in to group I and group II. To begin with, a detailed case history was taken for each participant to make sure that the participants have no symptoms of retro-cochlear pathology.Modified Hughston and Westlake method by Carhart and Jerger (1959), was used to measure the air conduction thresholds at octave frequencies from 250 Hz to 8 kHz and bone conduction thresholds for octave frequencies from 250 Hz 8 kHz. PTA_1(average of the pure tone thresholds at 500Hz, 1 kHz, 2 kHz) and PTA_2(average of the pure tone thresholds at 1 kHz, 2 kHz, 4 kHz) werecalculated to define the degree and configuration of hearing loss for group I and group II.

Speech identification score was obtained using the speech identification test material given by Yathiraj and Vijayalakshmi (2005), at a level of 40 dB SL (Re: SRT). Twenty five words were presented in the live mode (male speaker) at a level of 40 dBHL above PTA1. Scoring was done in such a way that each correct responses (repeating back correctly) got a score of 4%.Tympanometry was done using probe frequency of 226 Hz (Brooks, 1968; Holte, Margolis & Cavanaugh, 1991) at 85 dB SPL. Ipsilateral and contralateral reflexes were obtained

from test ear at 500 Hz, 1 kHz, and 2 kHz and for broad band noise. Mean threshold for group at each frequency is given in the Table 2.

Table 2: Mean thresholds at each audiometric frequency for each group

Degree of hearing loss	Duration of hearing loss (years)	Mean threshold (dBHL) across frequencies				
Minimal to moderate		***250 Hz***	***500Hz***	***1 kHz***	***2 kHz***	***4 kHz***
	< 2	18.33	22.55	37.55	48.33	60.00
	> 2	18.00	22.00	35.00	45.00	68.00
Moderate to severe	*< 2*	25.00	42.55	52.55	70.00	75.00
	> 2	42.88	50.77	60.00	70.00	80.71

Phase 2 - Selection of Hearing Aid

Four channel digital behind the ear hearing aid was selected to record aided ACC. The fitting range of the hearing aid covered severe to profound hearing loss. According to the manufacturer's specification, in 2cc coupler the frequency response of the hearing aid extended from 100 Hz to 6200Hz. Peak full on gain was 70dB and high frequency average full on gain was 62dB. Attack time was 10 ms and release time was 51 ms. The total harmonic distortion was less than 3% at 500 Hz, 2% at 800 Hz and 1% at 1600 Hz, whereas, the equivalent input noise was 18 dBSPL. The measured electroacoustic characteristics of the hearing aid were according to manufacturer's specification.

Phase 3 - Hearing Aid Fitting

After the selection of the hearing aid hearing aid fitting was done to the participants test ear. Regular ear mold that is attached to the BTE adapter was used to ensure that the squealing was absent and that is snugly fit into the participants test ear without creating discomfort. The following two approaches were made prior recording of ACC - i) programming of hearing aid ii) Real ear measurement of hearing aids.

Programming of Hearing Aid: The selected hearing aid was connected to the personal computer and Noah fitting software by using HiPro. The Bass Boost facility of the hearing aid was turned off, and the hearing thresholds were fed on to Noah fitting software and fitting module. The hearing aid was set in omni-directional mode with enabled compression circuits, where compression ratio in the NAL-NL1 default setting was used and volume control was disabled. NAL-NL1 fitting formula was used to prescribe the gain for hearing aid and hearing aid was programmed to match the target gain at acclimatization level two.

Real Ear Measurements: The participants were seated in front of the free field speaker of phoenix FP 7000 equipment. The speaker of equipment was placed at 45 degree azimuth and 1 meter distance from the participant. Then the sound field was equalized.

The audiogram of the test ear was plotted and fitting formula of NAL NL1, stimulus of digi-speech at 65 dB-SPL was selected in the phoenix FP 7000 software. The probe tube of the microphone was inserted by means of premeasured length, 25 to 30 mm past tragal notch of the test ear. Then, the hearing aid was fitted to the test ear using regular ear mold attached with a BTE adapter. The real gain was verified to match the NAL-NL1 target. The gain of the hearing aid was increased to match the target whenever the real ear SPL was not matching the target gain curve.

Phase 4 - Acquisition of Aided and Unaided Acoustic Change Complex

Once the gain of the hearing aid in the ear canal is verified through real ear measurement, ACC was acquired in aided and unaided condition in each participant. The participants were seated comfortably in an armed chair and the fitted hearing aid was switched to 'on' position. The electrode site was cleaned with skin preparation gel. Disc type gold coated electrodes was placed with the help of conduction gel at the test ear mastoid (M1/M2), upper forehead (Fpz), and vertex (Cz). The impedance was less than 5 kΩ and inter-electrode impedance within 2 kΩ. Caution was taken about the chance of closer proximity or contact of electrode and hearing aid microphone. The hearing aid was positioned to the periphery of the pinna and made sure that the hearing aid is not in closer proximity or contact of electrode and hearing aid microphone. For recording the ACC in aided conditions, the stimulus /si/ was presented through free field speaker, positioned at one meter distance at 45 ^{o} azimuth. The stimulus was presented at level of 65dBSPL and ACC was recorded for checking replicability of waveform. The non-test ear was blocked with ear mold impression material to avoid its participation (whenever required). The subjects were asked ignore the stimulus and watch a close captioned video while recording ACC.

After the acquisition of aided ACC the hearing aid was removed from participant's ear and unaided ACC was recorded. The non-test ear remained blocked with ear mold impression material to avoid its participation (whenever required).

Table 3: Stimulus and Acquisition parameters for recording ACC

Stimulus parameters	
Stimuli	/Si/
Duration of stimuli	250.8
Intensity	65dBSPL
Polarity	Alternating
Transducer	Loudspeaker
Mode of presentation	Free Field
Acquisition parameters	
Mode of stimulation	Ipsi
Electrode montage	Cz(+ve), M1/M2(-ve) of test ear and ground at Fz
Filter setting	0.1-30 Hz.
Analysis window	535 msec.
No. of channels	Single
Amplification	25,000
Repetition rate	1.1 per sec
Number of sweeps	150
No. of repetitions	2

Analysis of Waveforms

Analyses of unaided and aided ACC waveforms were done for all participants. The peak identification and morphology rating were done by two experienced audiologist in waveform analysis. The peaks which are marked by both audiologists were considered for analysis. When an audiologist marked a peak, but, not the other, the peak was not considered for analysis. The peaks of ACC were marked as P1, N1, P2, N2, P1', N1', P2', and N2'. Amplitude, latency, and morphology were the three measures considered for waveform analysis. The amplitude of the each peak was defined as the largest positive or negative deflection depending on whether it's a negative or positive peak in the response window. The latencies of the peaks were calculated by taking the center or midpoint when wave form contained double peak of equal amplitude. Latency was measure at the center of larger peak when peaks were not equal in amplitude (Ostroff, Martin &Boothroyd, 1998). The data point values representing N1, P2, N1' and P2' were tabulated for statistical analysis.

Results and Discussion

Three independent variables; subgroup, unaided, and aided condition were taken and their influence on the dependent variables; parameters of ACC (latency and amplitude) was studied. N1, P2, N1', P2' were the target parameters considered for statistical analysis.

Descriptive statistics (mean and standard deviation) were calculated for each subgroup; minimal to moderate sloping sensorineural hearing loss (SNHL) with duration of hearing loss less than two years (subgroup A), minimal to moderate sloping SNHL with duration of hearing loss greater than two years (subgroup B), moderate to severe sloping SNHL with duration of hearing loss less than two years (subgroup C), moderate to severe sloping SNHL with duration of hearing loss greater than two years (subgroup D).Mann Whitney test was administered to compare the unaided & aided ACC separately, between minimal to moderate & moderate to severe degrees of sloping sensorineural hearing loss, with duration of hearing loss less than and greater than two years.Mann Whitney test was administered to compare between less than or greater than two years of duration of hearing loss in minimal to moderate & moderate to severe degree of sloping sensorineural hearing loss separately.Wilcoxon signed rank test was done for comparing unaided and aided ACC within each subgroup. Figure 1 shows the unaided and aided waveforms of each subgroup.

Effect of Different Degree of Sloping SNHL on Unaided and Aided ACC

Mann Whitney test was administered to study the differential effect of different degree of sloping SNHL with duration of hearing loss less than two years on unaided and aided ACC. Latency and amplitude of N1' , P2' were compared separately between subgroup 'A' and 'C' in the unaided condition. N1, P2 peaks were absent in subgroup C, whereas, in subgroup A, the N1and P2 peaks were present only for two and three ears out of six ears. Hence, these ACC peaks were not included for comparison in the unaided condition. In the aided condition, N1 peak was present in all ears except one ear in the same subgroup. However, in subgroup C, it was present for all four ears. P2 peak was obtained for all ears in subgroup A. But, in subgroup C it was absent for one ear out of four ears aided condition. The reason for absence of N1 and P2 in subgroup C and its presence in subgroup A is discussed below.

The N1 of ACC represents the onset of consonant portion in the stimulus (syllable) (Ostroff, Martin & Boothroyd, 1998). The consonant portion (fricative portion) in the stimulus (/si/) has a spectral energy of 2 k Hz and above, more energy around 4 to 8 k Hz (shown in Figure 2). The audibility at the higher frequency is important for the perception of fricatives (Annabelle, Sher & Owens, 1972), and it is also evident from studies that if proper audibility is present there is indemnity of generation of far field recordable cortical potentials. In the present study, the two ears in the subgroup A had mean threshold at high frequency (from 1 k Hz to 4 k Hz) as 41.6 and 43.3 dBHL. Hence, the presence of N1 is accounted to the intensity level of the presentation of stimulus, which was at 65 dBSPL. But, in group C N1 was absent for all subjects. This suggests, audibility at

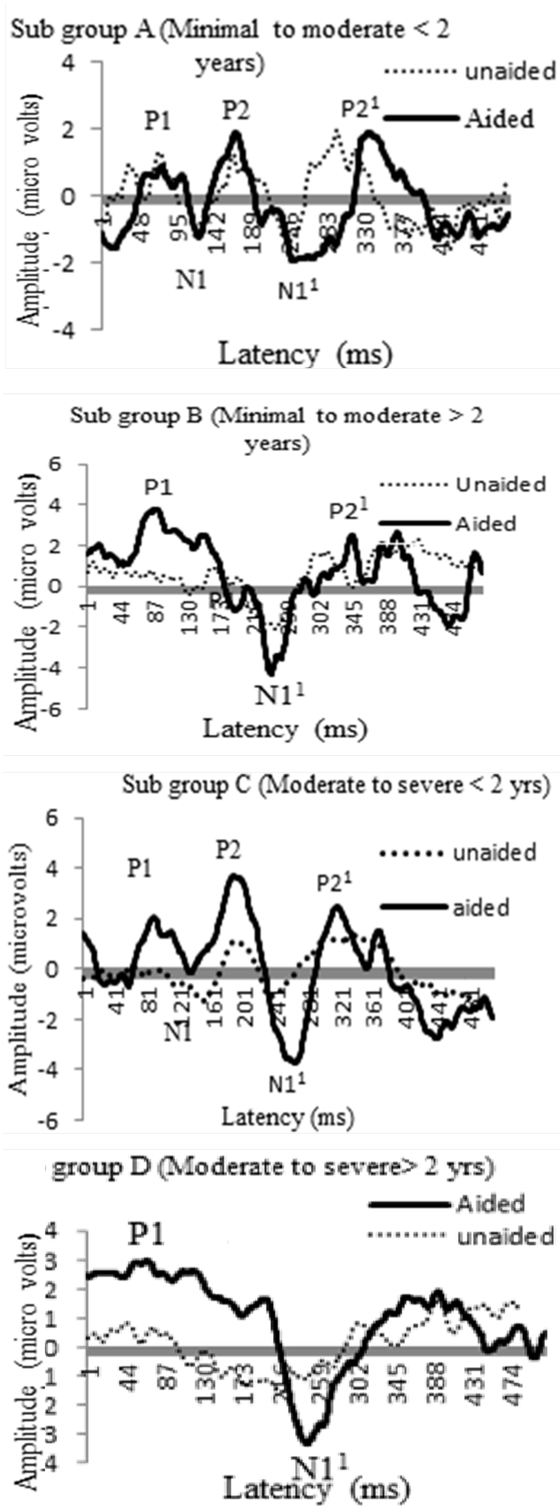


Figure 1: Unaided and aided waveforms for each subgroup.

the higher frequencies (1 k Hz and above) is important for the generation of N1.

Latency of N1', P2': The results of the Mann Whitney test for latency of N1' , P2' revealed that there is a significant difference in latency of P2' (Z = -2.07 p = .038) in unaided condition between subgroup A vs. C. But, aided condition did not show any statistical significance. The latency of P2'was shorter (mean latency of 321.93) for subgroup A compared to subgroup C (mean latency of 346.92 ms). There was no statistically significant difference in the N1'latency between subgroup A and C. The effect of degree of sloping SNHL on unaided and aided ACC is depicted in Figure 2 .

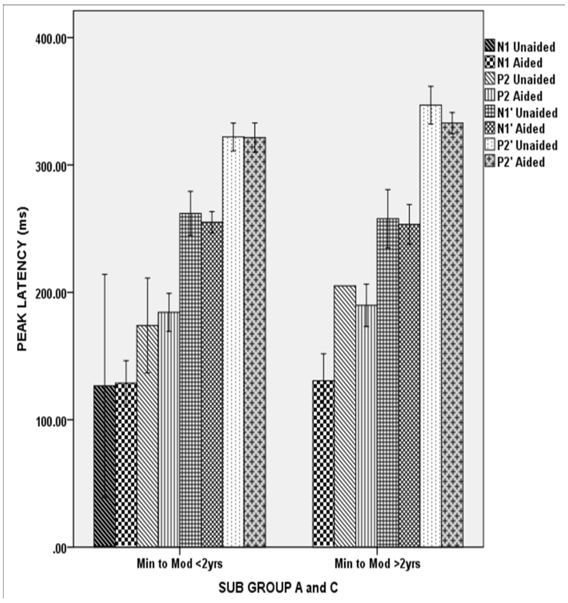


Figure 2: Effect of minimal to moderate to severe sloping SNHL with duration of hearing loss less than two years on latency of unaided and aided ACC.

Amplitude of N1', P2': Mann Whitney test revealed that there was no statistically significant difference in amplitude of P2 between subgroup A vs. C in both unaided and aided conditions. There was a statistically significant difference in amplitude of N1 (p = .020) in unaided condition between subgroup A vs. C, whereas, aided condition did not shown statistically significant differences. The amplitude of N1' was larger for subgroup A compare to subgroup C, individual with minimal to moderate sloping SNHL showed larger amplitude compared to moderate to severe sloping SNHL. The results are represented in Figure 3.

To conclude, there is a significant difference in P2' latency and N1' amplitude in unaided condition between minimal to moderate and moderate to severe sloping SNHL with durations of hearing loss less than two years. The P2' is a response to the vowel portion of the stimulus i.e. syllables (Ostroff, Martin, Boothroyd, 1998) and thresholds at the 250Hz to 2 kHz is important for the generation of P2'. In present study stimulus had a vowel (/i/) with more spectral energy in this region. Subgroup A had better thresholds at this frequency region compare to C as a result of this, the P2' latency was shorter for Subgroup A compare to C. Hence, it is possible to electrophysiologicaly state that when the degrees of slope till 2 kHz increases, the vowel coding at the cortical level decreases. The onset portion of /i/ vowel has got low frequency content than its steady

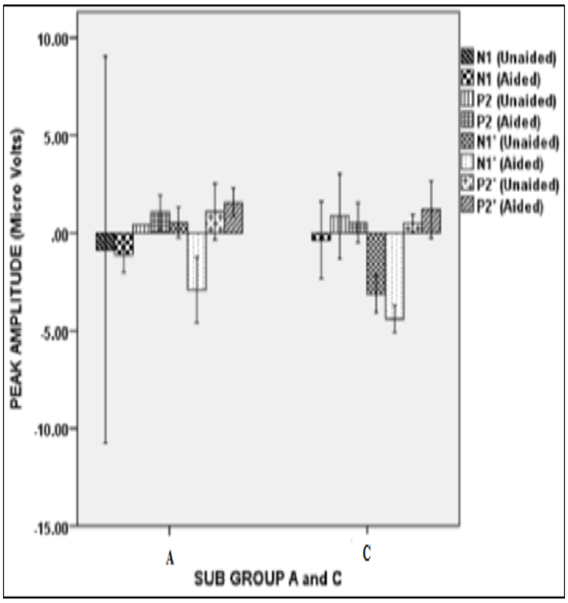


Figure 3: Effect of minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss less than two years on the amplitude of unaided and aided ACC.

state portion and cortical response to this onset portion is represented by N1' peak of ACC. The low frequency thresholds (250 Hz and 500 kHz - mean of 20.44 dBHL) were better for subgroup A than C (mean threshold of 33.77). So, there is larger amplitude of N1' for minimal to moderate than moderate to severe sloping SNHL with duration of hearing loss less than two years. But, in aided condition, there was no significant difference between two groups. It says that the hearing aid in both degrees (higher and lower) works similarly.

Effect of Minimal to Moderate vs. Moderate to Severe Sloping SNHL with Duration of Hearing Loss Greater than Two Years on Unaided and Aided ACC.

Latency of N1, P2, N1', P2': In unaided condition, N1 was absent for subgroup B and D, whereas, P2 was present only for two ear in the subgroup B with a mean latency of 203ms. Mann Whitney test revealed that there was no statistically significant difference in latency of N1, P2, between subgroup B vs. D in aided conditions. Latency of N1' (p = .059) in aided condition and P2' in unaided condition (p = 0.50) between subgroup B vs. D is approaching statistical significance. The latency of N1' in aided condition was longer (mean latency of 262.07) for subgroup B compare to subgroup D (mean latency of 252.78), i.e., minimal to moderate sloping SNHL group showed longer latency compare to moderate to severe sloping SNHL. The effect of minimal to moderate and moderate to severe SNHL on latency of unaided and aided ACC is depicted in Figure 4.

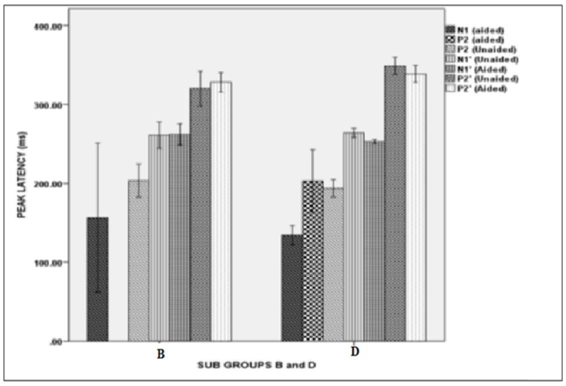


Figure 4: Effect of minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss more than two years on latency of unaided and aided ACC.

Amplitude of N1, P2, N1', P2': Mann Whitney test revealed that there was no statistically significant difference in amplitude of N1, P2, between subgroup B vs. D in aided conditions. In unaided condition, N1 was absent for subgroup B and D, and P2 was present only for two ear in the subgroup B with a mean amplitude of 2.67uv. Amplitude of N1'(p=0.053) and P2' (p=0.050) in unaided condition between subgroup B vs. D is approaching statistical significance. The amplitude of N1' in unaided condition was larger (mean amplitude of -3.9uv) for subgroup B compared to subgroup D (mean amplitude of -2.05μV), i.e., in individual with minimal to moderate sloping SNHL showed larger amplitude compare to moderate to severe sloping SNHL Effect of minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss more than two years on amplitude of unaided and aided ACC is represented in Figure 6.

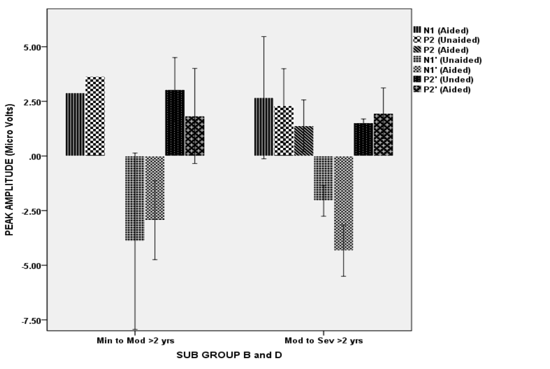


Figure 5: Effect of minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss more than two years on amplitude of unaided and aided ACC.

The comparison of greater duration (greater than two years) with two different degree of sloping SNHL demonstrated statistically different amplitude of N1' and P2'latency and amplitude in unaided condition. When, effect of greater duration of hearing loss added to degree of slope, N1' showed significant difference between minimal to moderate and moderate to severe sloping SNHL in unaided and aided condition. The auditory deprivation effect could be the probable reason for this. It indicates that if the duration of hearing loss is more, even with effective amplification also there is difference in cortical representation of speech sounds. The concept of auditory deprivation in individuals with moderate to severe SNHL as a measure of cortical potential was given by Buckley (2003). The result of the present study is in agreement with Buckley's study in terms changes in amplitude of the cortical potentials with auditory deprivation.

Effect of duration of hearing loss along with degree on unaided and aided ACC

Mann Whitney test was administered to study the effect of duration of hearing loss on a particular degree of sloping hearing loss. The results of effect of duration of hearing loss on degree of sloping SNHL in unaided and aided ACC conditions are discussed in the following section

Effect of duration of hearing loss on minimal to moderate sloping SNHL

Mann-Whitney test was administered to study the two different duration of hearing loss; less than two years and greater than two years on minimal to moderate sloping SNHL. Latency and amplitude of N1, P2, *N1',P2'*in subgroup A is compared with the same of subgroup B to study the effect of duration of hearing loss on ears with minimal to moderate sloping SNHL.

Latency of N1, P2, N1', P2': Mann - Whitney test revealed that there was no statistically significant difference in latencies of N1, P2, N1', P2' between subgroup A vs. B in unaided condition. The N1 is present only for two ears in subgroup A (mean latency of 126.61ms) out of six ears, whereas, it was absent for all ears in subgroup B. P2, is present in three out six ears (mean latency of 174.01)in subgroup A and two ears out of four ears (mean latency of 203ms) in the subgroup B. But, latency of P2in aided condition is approaching statistical significance (p = 0.052) between subgroup A vs. B. The latency of P2was shorter (mean latency of 184.30ms) for subgroup A compared to subgroup B (mean latency of 203.26ms).

Amplitude of N1, P2, N1', P2': Mann Whitney test revealed that there was no statistically significant difference in amplitude of N1, P2, between subgroup A vs. B in aided conditions, but these peaks were not considered in the unaided condition as N1 was absent for all ears in subgroup B and P2 was only present for three (mean amplitude of 0.44μV) and two ears (mean amplitude of 2.67μV) in subgroup A & B respectively. Amplitude of N1' in unaided condition between subgroup A vs.

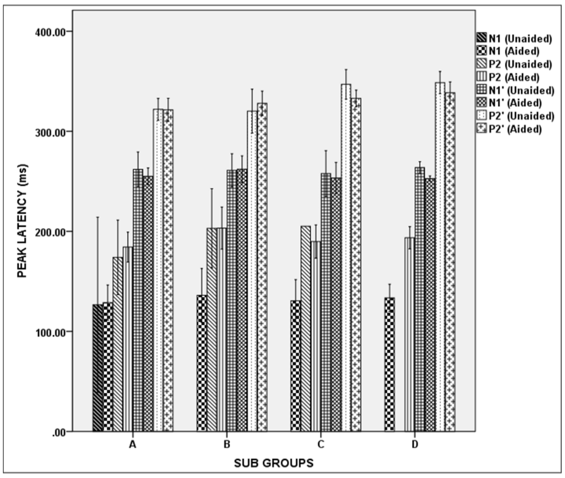


Figure 6: Effect of duration of hearing loss (less than two years and more than two years) of minimal to moderate and moderate to severe sloping SNHL on latency of unaided and aided ACC.

B is statistical significant (p = 0.020). The amplitude of N1'in unaided condition was larger for subgroup A compare to subgroup B.

To conclude, minimal to moderate sloping SNHL with different durations are not affecting the latency of unaided ACC, but, the amplitude of N1'. This effect was more seen in minimal to moderate sloping SNHL with duration of hearing loss greater than two years. The probable reason is auditory deprivation. In aided condition, P2 showed a difference, this tells us that with amplification there is difference in consonant coding at cortical level in individual with more than two years of duration of minimal to moderate sloping SNHL. So, it could be that the auditory deprivation within minimal to moderate sloping SNHL has an effect on cortical neural coding of acoustical changes within speech sounds with amplification.

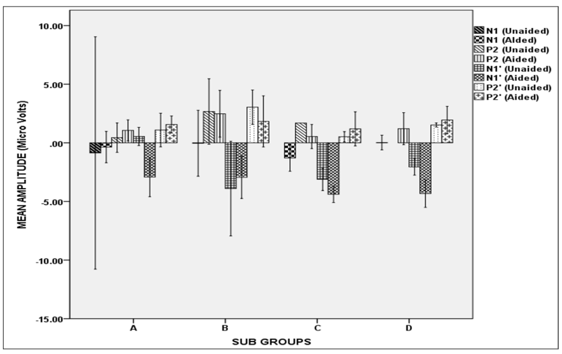


Figure 7: Effect of duration of hearing loss (less than two years and more than two years) of minimal to moderate and moderate to severe sloping SNHL on amplitude of unaided and aided ACC.

Effect of duration of hearing loss on moderate to severe sloping SNHL

Latency and amplitude of N1, P2, N1', P2' in subgroup C is compared with D to study the effect of duration of hearing loss on ears with moderate to severe sloping SNHL.N1 and P2 were absent in subgroup C & D in unaided condition, hence these peaks were not taken for comparison in unaided condition. But, N1, P2, N1'and P2'peaks were considered for comparison in aided condition.

Latency of N1, P2, N1', P2': In unaided condition, Mann Whitney test revealed that there was no statistically significant difference in latencies of N1' and P2' between subgroup C and D, whereas, in aided condition latencies of N1, P2, N1',P2' also does not shown any statistically significant differences. Effect of duration of hearing loss on latency of unaided and aided ACC is represented in figure7.

Amplitude of N1, P2, N1', P2': Mann Whitney test revealed that there is a statistically significant difference in amplitudes of P2' (p = 0.034), between subgroup C vs. D. Whereas, N1' is approaching statistical significance (0.059) in unaided condition. In aided condition, only the amplitude of showed statistically significant difference (p = 0.019) between subgroup C &D. Effect of duration of hearing loss on amplitude of unaided and aided ACC is represented in Figure 8.

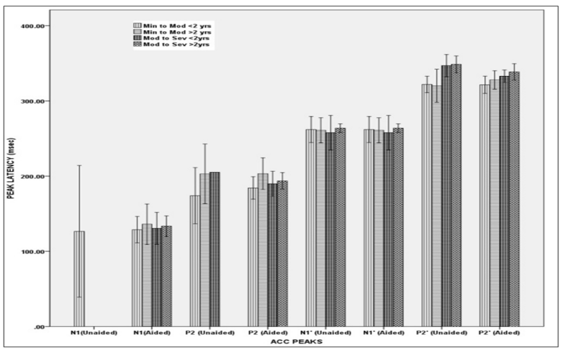


Figure 8: The effect of different duration of hearing loss on latency of each ACC peaks in unaided and aided condition.

The individual with moderate to severe sloping SNHL with duration of hearing loss less than two years or greater than two years had no N1 and P2 peaks of ACC. This indicates that consonant (fricative /s/) is not coded at their auditory cortical structures which are responsible for the generation of ACC potential at usual conversational speech level (65 dBSPL). The reason for this is poor thresholds at the higher frequencies (frequencies above 2 KHz) leading to lack of 'audibility' of high frequency, low intensity speech sounds like fricatives. Lack of audibility at high frequencies could not be the only reason even though it is the major one. The other probable reason could be that the auditory deprivation may induce changes in the amplitude and latency of ACC, which may act together with decreased audibility at an input level of 65 dBHL. Hence, to support this, further research with various input levels to the hearing aids, needs to be conducted which will give information about auditory deprivation of ACC, as at higher levels the effect of audibility is compensated.

However, in these individuals there was a presence of N1' and P2' peaks, which indicate audibility to the vowel portion, can be preserved to some extent in these individuals with sloping SNHL. But, there is a question which need to be answered is, why the vowel coding was not significantly different (no significant difference in latency and amplitude N1' and P2') between duration of hearing loss of less than two years and more than two years. The probable reason for this is the heterogenic cortical auditory response characteristics in individual with moderate to severe sloping SNHL, and also the variability in the participant selection (number of ears selected in subgroup C had only four, which, D had only seven). However, a proper amplification is given to these participants; there was a presence of N1, P2 peaks and statistically significant difference in N1' amplitude of aided ACC between moderate to severe sloping SNHL with duration of hearing loss less than two years (more amplitude)and more than two years (less amplitude). The reason for presence of N1, P2 indicates that with proper amplification the consonant (/s/) is coded in the cortical level of individual with moderate to severe sloping SNHL irrespective of duration of hearing loss. But, the reason for significant difference in N1' amplitude (indicates vowel onset portion (/i/) in the stimulus coded differently in moderate to severe sloping SNHL with duration of hearing loss less than two years and greater than two years is probably auditory deprivation.

Only N1'showed significant difference because this response is mainly elicited by the low frequency portions of vowel (/i/) and it is already evident from electrophysiological studies that cortical response to the low frequency stimulus is better than response to the high frequency stimulus (Picton, Woods &Proulex, 1978). The reason is that, inherent lower amplitude of the cortical response when elicited from stimulus containing high frequency content such as /ʃ/ and /s/ compared to /m/, /a/, /u/ and /i/ which had predominant low frequency content as documented by Agung, Purdy, McMahon &Newall (2006). One of the stimulus related reason is that /i/ portion of /si/ stimulus used in the present study had larger amplitude compared to /s/ portion. Hence, the better amplitude of N1'.

To explain further, physiologically, individuals with cochlear pathology have wider auditory filter, is a reason they are more prone to the effects of upward spread of masking (Moore, 1998). The preceding high fre-

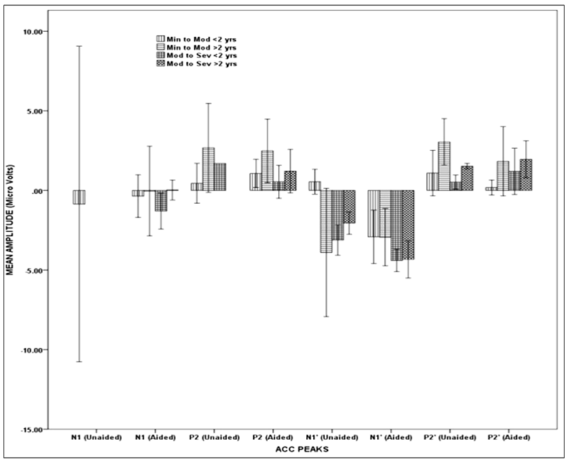


Figure 9: Effect of duration of hearing loss on amplitude of unaided and aided ACC.

quency signals (fricative /s/) in the stimulus of this study could have been masked to an extent by low frequency vowel portion. It is evident from studies that low frequency portions of speech like vowels may mask higher frequency components like frication noise as reported by Dillon (2001). Hence, the N1' amplitude was higher compare to other peaks.

Comparison of Aided and Unaided ACC within Each Subgroup

Latencies and amplitude of P2, N1', $P2^1$ (compared only within subgroup A) and N1', P2' (compared within subgroup B, C, D) were compared using Wilcoxon Signed Ranks test between unaided and aided condition within each subgroup. Latency and amplitude comparison of unaided and aided peaks of ACC were done, in order to study the effect of amplification along with degree and different duration of hearing loss on cortical neural representation of acoustical changes within speech sounds (syllable).

Comparison of aided and unaided ACC within subgroup A

The latency and amplitude of P2, N1', and P2'peaks were compared separately between unaided and aided condition within subgroup A using Wilcoxon Signed Ranks test. The results of wilcoxon Signed Rank test for each subgroup is shown in Table 3.

Latency of P2, N1', P2': The results of the test revealed that there are no statistically significant difference latencies of P2, N1', and $P2^1$ between unaided and aided condition within minimal to moderate sloping SNHL with duration of hearing loss less than two years.

Amplitude of P2, $P2^1$ also does not shown any statistically significant differences, whereas, N1' has shown a statistically significant difference ($p = 0.028$). The amplitude of (mean amplitude of 0.54, shown in table 4.1)

Table 4: Results of Wilcoxon Signed Ranks Test

Subgroups	Comparisons	Z value	p
A	*Latency of P2 (aided) - latency of P2 (unaided)*	-0.535	*0.590*
	Latency of N1' (aided) - latency of N1' (unaided)	-1.15	0.249
	Latency of P2' (aided) - latency of P2' (unaided)	*-0.420*	0.674
	Amplitude of P2 (aided) -amplitude of P2 (unaided)	-0.535	0.593
	Amplitude of N1' (aided) - Amplitude of N1' (unaided)	-2.20	0.028*
	Amplitude of P2' (aided) - Amplitude of P2' (unaided)	-1.15	0.25
B	*Latency of N1' (aided) - latency of N1' (unaided)*	0.000	1.00
	Latency of P2' (aided) - latency of P2' (unaided)	-1.06	0.28
	Amplitude of N1' (aided) - Amplitude of N1' (unaided)	-1.06	0.10
	Amplitude of P2' (aided) - Amplitude of P2' (unaided)	-0.535	0.59
C	*Latency of N1' (aided) - latency of N1' (unaided)*	-0.730	0.46
	Latency of P2' (aided) - latency of P2' (unaided)	-1.82	0.06
	Amplitude of N1' (aided) - Amplitude of N1' (unaided)	-1.82	0.06
	Amplitude of P2' (aided) - Amplitude P2' (unaided)	-1.46	0.14
D	*Latency of N1'(aided) - latency of N1' (unaided)*	-2.19	0.028*
	Latency of P2' (aided) - latency of P2' (unaided)	-1.60	0.10
	Amplitude of N1' (aided) - Amplitude of N1' (unaided)	-2.36	0.01*
	Amplitude of P2' (aided) - Amplitude of P2' (unaided)	-1.60	0.10

*Indicates statistically significant. *Amplitude of P2, N1', P2'*

N1' in aided condition is larger than in unaided (mean amplitude of -2.91, shown in table 4) within subgroup A.

Comparison of aided and unaided ACC within subgroup B and C

The latency and amplitude of N1', P2[1]peaks were compared separately between unaided and aided condition within subgroup B and C using Wilcoxon Signed Ranks test.

Latency of N1', P2': The results of the test revealed that there are no statistically significant difference in the latencies of N1', P2[1] between unaided and aided condition within minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss greater than and less than two years.

Amplitude of N1', P2': Amplitude of N1' and P2[1]did not show any statistically significant differences between unaided and aided condition within minimal to moderate and moderate to severe sloping SNHL with duration of hearing loss greater than & less than two years respectively.

Comparison of aided and unaided ACC within subgroup D

The latency and amplitude of N1', P2[1]peaks were compared separately between unaided and aided condition within subgroup D using Wilcoxon Signed Ranks test.

Latency of N1', P2': The results of the test revealed that there are no statistically significant difference in the latencies of P2[1]but, there is a statistically significant difference in N1' latency (p = 0.028) between unaided and aided condition within moderate to severe sloping SNHL with duration of hearing loss greater than two years .

Amplitude of N1', P2': Amplitude of N1' is also shown statistically significant differences (p = 0.018) between unaided and aided condition within moderate to severe sloping SNHL with duration of hearing loss greater than two years. But, there is no statistically significant difference in amplitude of P2'.

Justification for the results under within subgroup comparisons

The results of the within group comparison of unaided and aided ACC in subgroup A, clearly throws light on to the effectiveness of amplification in sloping hearing loss with duration of hearing loss less than two years. There was appearance of N1 in aided condition; which indicates that the audibility of (Korczak, Kurtzberg&Stapells, 2005; Oates, Kurtzberg&Stapells, 2000;Polen, 1984), the fricative portion was enhanced by amplification. The reason for absence of N1 in subjects minimal to moderate sloping hearing loss can be the upward spread of masking (Moore, 1998). In the present study low frequency - high amplitude vowel could have been masked the onset portion of fricative /s/. So, even minimal to moderate sloping SNHL with duration of hearing loss less than two years needs to be provided with hearing aids, even if they have minimal speech perception problems. The above statement has to be supported by further behavioral and electrophysiological studies.

The results of the within group comparison of unaided and aided ACC in subgroup B, throws light on to the effectiveness of amplification in sloping hearing loss with duration of hearing loss more than two years. It is indicated that latency and amplitudes of unaided and aided condition did not show any significant differences, even though there is a difference in the mean value. The probable reason for this is auditory deprivation.

The results of the within group comparison of unaided and aided ACC in subgroup C, gives information on the effectiveness of amplification in moderate to severe sloping hearing loss with duration of hearing loss less than two years. There was appearance of N1, P2 in aided condition because of audibility at the higher frequency provided by the hearing aid. It is indicated that latency and amplitudes of unaided and aided ACC did not show any significant differences, though there is a difference in the mean value. The probable reason could be the abnormal growth of loudness phenomenon that occurred at the threshold level (Florentine, Fastle&Buus, 1998) in these participants in the unaided condition but not in aided condition with probably may be contributed to the amplification factors by the hearing aid.

The results of the within group comparison of unaided and aided ACC in subgroup D, clearly throws light on to the effectiveness of amplification in moderate to severe sloping hearing loss with a duration of hearing loss more than two years. There was appearance of N1, P2 peak in aided condition even in individual with more than two years of deprivation because of audibility at the higher frequency provided by the hearing aid. It is indicated that latency and amplitudes of unaided ACC peaks (N1', P2') did not show any significant differences but there was a difference in latency and amplitude of N1 ' of aided ACC compare to unaided ACC. The reason for this finding is presently unclear.

Conclusions

From the study we can conclude that since most of the peaks of ACC are present in aided condition compared to unaided condition in individual with varying degrees of sloping sensorineural hearing loss, ACC can be used as an objective measure to quantify the benefit from amplification.Vowel portion of the speech stimulus is better coded in the cortical level compared to consonant portion as evidenced from larger amplitude of N1' and P2' compared to N1 and P2 in individuals with varying degrees of sloping hearing loss.It is evident from the results of this study that the hearing aids are not able to give adequate amplification at higher frequencies as the fricatives are not being coded well at the cortical level compared to vowels. Hence, measures should be taken to improve the spectral content of the higher frequency sounds by the amplification device for individuals with sloping sensorineural hearing loss.

References

Agung, K., Purdy, S. C., McMahon, C. M., & Newall, P. (2006). The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *Journal of the American Academy of Audiology, 17*(8), 559-572.

Brooks, D. N., (1998). An objective method of determining fluid in the middle ear.*International Audiology, 7,* 280-286.

Buckley, K. A. (2003). Effect of auditory deprivation in cortical processing.*The Hearing Journal, 56*(6), 16-18.

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

Coughlin, M., Kewley-Port., & Humes. (1998). The relation between identification and discrimination of vowels in young and elderly listeners. *Journal of the Acoustical Society of America, 104*(6), 3457-64.

Dillon, H. (2001). Hearing aids. Turramurra, Australia: Boomerang Press.

Dubno, J. R., Dirks, D. D., & Schaefer, A. B. (1987).Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *Journal of the Acoustical Society of America, 81,* 1940-1947.

Florentine, M., Fastl, H., & Buus, S. (1988). Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking.*Journal of the Acoustical Society of America, 84*, 195-203.

Holte, L., Margolis, R. H., & Cavanaugh, R. M. (1991). Developmental changes in multifrequency tympanometry. *Audiology, 300,* 1-24.

Kartik, N., & Vanaja, C. S. (2005).*Acoustic Change Complex ACC : An Electrophysiological index for Speech Perception in Children and Adults.* Unpublished Masters's Dissertation. University of Mysore, Mysore.

Korczak, P. A., Kurtzberg, D., & Stapells, D. R. (2005).Effects of sensorineural hearing loss and personal hearing aids on cortical event-related potential and behavioural measures of speech-sound processing.*Ear and Hearing, 26,* 165-185.

Martin, B. A., & Boothroyd, A. (1999).Cortical auditory event-related potentials in response to periodic and aperiodic stimuli with the same spectral envelope.*Ear and Hearing, 20*(1),33-44.

Moore, B. C. J., (1998) *Cochlear hearing loss*. London: Whurr Publishers Ltd.

Oates, P. A., Kurtzberg, D., & Stapells, D. R., (2002). Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing.*Ear and*

Hearing, 23, 399-415.

Ostroff, J. M., Martin, B. A., & Boothroyd, A. (1998). Cortical evoked response to spectral changes within a syllable. *Ear and Hearing 19*(4), 290-297.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978). Human auditory sustained potentials: II. Stimulus relationships.*Electroencephalography and Clinical Neurophysiology, 45*, 198-210.

Pittman, A. L., &Stelmachowicz, P. G. (2003). Hearing loss in children and adults: audiometric configuration, asymmetry, and progression. *Ear and Hearing, 24,* 198-205.

Polen, S. B. (1984). Auditory event related potentials. *Seminars in Hearing, 5,* 127-141.

Sher, A. E., & Owens, E. (1974). Consonant confusion associated with hearing loss above 2 kHz. *Journal of Speech and Hearing Research, 17,* 669-681.

Silman, S., Gelfand, S., & Silverman, C. (1984). Late onset auditory deprivation: effect of monaural versus binaural hearing aids. *Journal of the Acoustical Society of America 76*(5), 1357-1362.

Tremblay, K. L., Friesen, L. M., Martin, B. A., & Wright, R. (2003). Test-retest reliability of cortical evoked potentials using naturally-produced speech sounds. *Ear and Hearing, 24*, 225-232.

Tremblay, K. L., Kalstein, L., Billings, C. J., & Souza, P. E. (2006).The neural representation of consonant vowel transitions in adults who wear hearing aids.*Trends in Amplification, 10,* 155-162.

Yathiraj, A., & Vijayalakshmi, C. S. (2005).*Phonemically Balanced Word List in Kannada.* A test developed in Department of Audiology, AIISH. Mysore.

Mizo High Frequency Speech Identification Test (MHF-SIT)

[1]Jonathan Lalchawilena & [2]Chandni Jain

Abstract

The aim of the present study was to develop a Mizo High Frequency Speech Identification Test (MHF-SIT) for the native speakers of Mizo having a high frequency hearing loss. The study also aimed to normalize and standardize the developed test on 100 adults with normal hearing ability. The developed test material consisted of Monosyllabic and Bisyllabic word lists which were further divided into 6 half lists containing 25 words each. Results revealed that there was no significant difference in speech identification scores between the 6 half lists. There was also no significant difference between the performances of male and female on the 6 half lists when the statistical test was carried out. The results demonstrated that the developed test material can be used to measure speech identification scores for persons having a high frequency hearing loss.

Keywords:*High Frequency word list, Sloping hearing loss, speech identification scores.*

Introduction

Hearing is the ability to perceive sounds by detecting vibrations through the organs of hearing. The basic mechanism for speech perception and effective communication precludes the ability to hear speech accurately. Speech perception involves the process by which the sounds of language are heard, interpreted and understood. The speech signal contains a number of acoustic cues like voice onset time, place of articulation or manner of articulation that are used in speech perception. These acoustic cues and other phonetic information are used for higher language processing and word recognition. An individual needs good hearing ability in order to decode a message from a stream of sounds coming from a speaker for effective communication (Borden & Harris, 1980) and it is a known fact that a person with hearing loss will have a communication problem.

In acoustic terms, vowels and consonants are described by their average pitch (frequency, measured in Hertz - Hz) and their average loudness (intensity, measured in decibels - dB) conversational speech level. The conversational speech has the most acoustic energy between 500 Hz and 3000 Hz. This frequency region is important for understanding meaningful speech (Pavlovic, 1987; Studebaker & Sherbecoe, 2002). A person needs to hear up to 1000Hz in order to hear all the vowels, but he/she needs to hear up to 3000 Hz in order to discriminate between the vowels and this helps us to discriminate between words that otherwise sound same.

However, the speech energy above 3000 Hz is also reported to offer listeners important linguistic information. Thus, one needs to have normal hearing sensitivity at high frequencies for good speech recognition. Any pathology that reduces hearing sensitivity at high frequencies will also affect speech recognition. Therefore, it is the essential duty of audiologists to identify, evaluate and rehabilitate aurally handicapped individuals.

Pure tone audiometry is a hearing test that involves the presentation of a series of sinusoidal tones (or beeps), to the listener, at specific frequencies to establish a person's hearing acuity. However, pure tone audiometry can only assess the auditory system's ability to "carry a simple stimulus" (Egan, 1979), and provides little information about the individual's ability to understand speech; speech audiometry is a type of audiological evaluation designed to better describe communication abilities and attempts to indicate how an individual may hear and understand spoken discourse by presenting words, instead of tones, during testing (Epstein, 1978). Also, because speech is a complex and continually varying signal requiring multiple auditory discrimination skills, it is not possible to accurately predict an individual's speech recognition from the pure-tone audiogram (Marshall & Bacon, 1981). Thus, pure tone audiometry provides little information about the individual's ability to understand speech.

Speech audiometry is routinely administered by audiologists as part of their audiologic evaluation battery. Basic speech audiometric tests include speech detection threshold (SDT), speech recognition threshold (SRT), speech identification scores (SIS), most comfortable level (MCL), and uncomfortable level (UCL). The purpose of these tests is to validate the pure tone air-conduction threshold and provide an index of hearing sensitivity for speech (Carhart, 1952; Chaiklin & Ventry, 1964). Speech audiometry materials were first developed in Standard American English but are now available in various other languages such as Hindi (Abrol, 1971; De, 1973), Spanish (Harris & Christensen, 1996), Italian (Turrini et al.,1993; Greer, 1997), Portuguese (Harris, Goffi, Pedalini, Merrill, & Gygi, 2001, Polish (Harris, Nielson, McPherson, & Skarzynski, 2004), Mandarin Chinese (Nissen, Harris, Jennings, Eggett, & Buck, 2005), Russian (Harris, Nissen, Pola, McPherson, Tavartkiladze & Eggett, 2007), Tongan (Seaver, 2008), etc. However, all these materials are conventional or regular speech audiometry materials

[1]Email: thantea2007@rediffmail.com
[2]Lecturer in Audiology, Email: chandni_j_2002@yahoo.co.in

which were developed and standardized on individuals who have mostly flat or nearly flat type of audiometric configuration. The use of these standard speech tests can give a prediction of the best hearing threshold levels in the mid frequency region of the auditory range.

It is also a known fact that a person with hearing loss is bound to have difficulty in perceiving speech. The kind and degree of perceptual difficulty depends on several factors, which include the degree of hearing loss, the type of hearing loss and the configuration of the audiogram (Gardner, 1971; Jerger & Jerger, 1971; Pascoe, 1975; Owens & Schubert, 1977; Lacroix & Harris, 1979). Depending on the audiometric configuration the speech perception ability would vary. A person with a high frequency hearing loss would have difficulty, mainly in hearing, speech sounds having an energy concentration in the high frequency regions (Stark, 1979; McDermontt & Dean, 2000). Martin (1987) concluded in his study by saying that speech perception would vary depending on whether the person had gradually sloping, sharply sloping or precipitously sloping audiograms. Another study by Mascarenhas (2002) also indicated that individuals with a sharply sloping high frequency hearing loss perform poorer than those with gradual and precipitous sloping high frequency hearing loss on High Frequency-Kannada Speech Identification Test (HF-KSIT).

Speech is a redundant stimulus as it contains information that is conveyed in various ways simultaneously (Martin, 1994). A hearing loss involving only part of the auditory frequency range may go undetected in a speech test if it is not carefully controlled. Thus, the use of a regular speech identification test would be insensitive towards identification of the problem of a person with a sloping high frequency hearing loss. In regular speech material, low frequency information may contribute redundant cues to the perceptual ability, thereby decreasing the sensitivity of the test in detecting their communication handicap (Schwartz & Surr, 1979; Kiukaanniemi & Maatta, 1980).

Owen & Schubert (1977) developed 100 multiple choice items for consonant identification in the California Consonant Test (CCT) to use with hearing impaired patients. A computer assisted analysis was obtained from the test responses of 550 patients with sensorineural hearing loss. They found that the test seems highly sensitive to configurations of high-tone loss, but were not sensitive to a regular speech test, i.e. CID W-22. They concluded that the two test measures different aspects of speech perception and the regular speech test does not assess the real communication problem of individuals with a high frequency loss. Similar results were obtained by Chung and Mack (1979) indicating that in quiet condition, both normal hearing participants and individuals with a high frequency hearing loss performed equally on a regular speech test (i.e. CID W-22). This indicated that the test was not sensitive to the communication problems of individuals with high frequency sensorineural hearing loss.

Also, selection of profitable hearing aids for elderly people or who have sloping high frequency hearing loss depends on utilizing a test which is sensitive to their problem (Sudipta, 2006). A significant improvement in speech identification scores was reported between the aided and unaided scores when High Frequency-Kannada Speech Identification Test (HF-KSIT) was administered instead of Common Speech Identification Test (CSDTI) developed by Mayadevi (1974) in individuals with high-tone loss (Mascarenhas, 2002). Thus, the use of speech identification test normally used may not determine their true communication handicap giving a maximum score unaided. It is unlikely that a person with sloping high frequency hearing loss will get maximum score unaided if a test material is used which are specially designed for their type of hearing loss.

Thus, it can be concluded that, regular speech identification tests are not sensitive to assess the perceptual problems of individuals with a sloping high frequency hearing loss. Hence, there is a need for special test to be developed and used while testing them. Such special tests have been developed in the past and they are called High Frequency Speech Identification Test. First high frequency word list was developed by Gardner in Standard American English (Gardner, 1971). The test contains consonants of high frequency spectral energy and is used for testing speech identification in cases of high frequency hearing loss. Currently, knowing its importance, these kind of tests are also developed in different Indian languages such as High frequency speech identification test in Hindi (Ramachandra, 2001), High frequency-Kannada speech identification tests (Mascarenhas, 2002), **High Frequency-English Speech Identification Test (HF-ESIT)** (Sudipta, 2006), High Frequency Speech Identification test in Tamil (Sinthiya, 2009), and High Frequency Speech Identification test in Telegu (Ratnakar, 2010). However, such materials are not currently available in the Mizo language. The purpose of this study was therefore to develop high frequency speech identification tests materials for native speakers of Mizo.

Method

The study was conducted in the two phases; the first phase involved the development of high frequency word list and the second phase involved standardization of the developed test material.

Phase I -The Development of High Frequency Word List

Both monosyllabic and bisyllabic words with good redundant cues were selected for the construction of the test list. These words were collected from various sources like dictionaries, newspapers, articles, and books. Words having a phoneme /k/, /t̪/, /s/, /d̪/, /r/, /dʒ/, and /l/ were preferred for inclusion as these phonemes have spectral energy mostly distributed above 1000 Hz frequency (Hughes & Halle, 1956). From vowels, the vowels like /i/ and /e/ were preferred as these vowels show higher F2 and F3 formants (Copper, Liebermann, Delattre, Borst & Gerstmann, 1952). Totally, 445 words were collected from the various sources mentioned.

To assess the familiarity of the words selected, 10 adults (5male & 5female), who were native speakers of Mizo were chosen and instructed to rate the selected words according to their level of familiarity. A three point scale of familiarity was administered based on the frequency of occurrence: most familiar, familiar and unfamiliar. The operational definition of most familiar words are those words used commonly, familiar words are those words used occasionally and unfamiliar words are those words that are not used by the participants.

A printed version of words was given to the individuals for the assessment with instruction carefully written to follow while rating each word. Each participant was asked to tick the score suitable. Only those words that were rated as most familiar were selected for the construction of test lists. Thus, 378 words were selected out of 445 words.

After the assessment of familiarity, LTASS was done to determine if the most familiar words constitute spectral information predominantly in the higher frequencies. This was done to see if the selected word lists possess the required spectral information. It was a necessary procedure because the spectral information of the phonemes /k/, /t̪/, /s/, /d̪/, /r/, /dʒ/, and /l/ could differ depending on the context and hence the so called high frequency word may not be having high frequency spectral information. To do this, 378 words that were

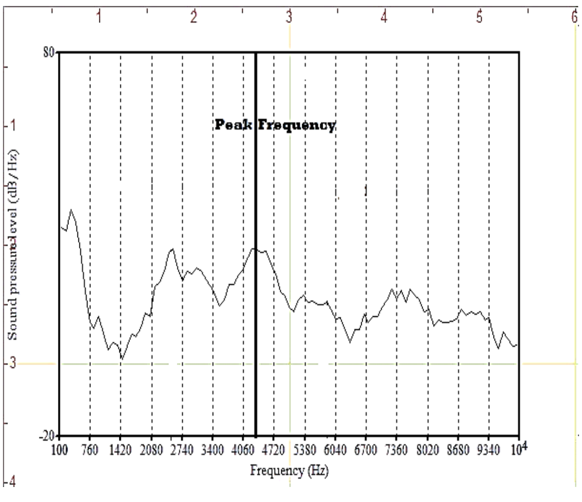


Figure 1: Spectrum derived from LTASS for a representative word (e.g. beisei).

rated as most familiar were assessed for the spectral information using LTASS. These words were recorded in a sound treated room using Computer software. They were spoken by adult female who was a native speaker of Mizo. The recording was done at 16 kHz sampling rate and 16 bit quantization using Adobe Audition software (version 3.0) and the samples were stored into a computer. LTASS was derived using the PRAAT software by feeding the audio samples one by one into it and the spectral information was determined manually. The peak frequency of the spectra which have a higher energy concentration was taken as the target parameter.

Peak frequency determined in LTASS showed that out of 378 words, there were 250 words having highest energy above 1000 Hz. These 250 words were further categorized based on different cutoff peak frequency (1.0, 1.5, 2.0, 2.5, & 3.0 kHz). In the present study it was decided to use words with peak frequency 2 kHz and above, to make it a more sensitive test. There were 186 such words with peak frequency of 2 kHz and above. Out of 186 words, 150 words were randomly selected for the construction of high frequency word list.

The spectrum derived from LTASS for a word is shown in the Figure 1. The 150 words that were selected based on LTASS results and having the highest energy above 2 kHz were further categorized into monosyllabic and bisyllabic words. There were 100 monosyllabic words and 50 bisyllabic words. The 100 monosyllabic words were divided into two lists, each of which were further divided into two half lists, each list containing 25 words. There were 50 bisyllabic words which were also divided into two half lists, each list containing 25 words. The frequency of occurrence of high frequency sounds was maintained same in each of the lists, containing both monosyllabic and bisyllabic words.

Recording of the test material (150 selected words) was done in a sound treated room as per ANSI guidelines (1999). Three adult females, who were native speakers of Mizo was selected for recording the tests words. The recording was done using Adobe Audition (version 3.0) software and the recorded materials were perceptually rated. The speaker who spoke with the best clarity and fluency was chosen for the audio recording of the final test list. The microphone was placed at a distance of 5 inches away from the mouth of the speaker. The VU meter was monitored within optimum levels during the recording. The speaker was instructed to say the words with her normal pitch and to keep the loudness constant across the words. The recording was done using 44.1 kHz sampling rate and 16 bit quantization in mono channel. The intensity of each word was edited as a single utterance using Adobe Audition software to obtain the same average RMS power as a 1000Hz calibration tone in an attempt to equate test word audibility (Harris et al., 2004; Wilson & Strouse, 1999). After editing, each word was saved individually as wave file.

The recorded material was also edited to carry out noise and hiss reduction. The inter stimulus interval between the two words was set to 5 seconds. The material was then copied onto an audio compact disc using a compact disc writer.

Phase II - Standardization of the Test Material

Participants:

The developed test MHF-SIT was normalized and standardized by obtaining speech identification scores on 100 native speakers of the Mizo language (50male & 50female). The participants were in the age range of 18 to 30 years. By self-report, all participants were native speakers of Mizo and considered Mizo to be their first language.

The participants of the study had normal hearing, as indicated by their four-frequency (500 Hz, 1000 Hz, 2000 Hz & 4000 Hz) pure-tone average threshold of ?15 dB HL, 'A' Type tympanogram with acoustic reflex thresholds in normal limits (90 dB at 1000 Hz). It was ascertained from a structured interview that none of these listeners had any difficulty in understanding speech in daily listening conditions and that they did not have any history of neurologic or otologic disorder.

Instrumentation:

All the evaluations were carried out in an acoustically treated two-room situation as per ANSI S3.1 (1991) using a calibrated single channel audiometer (Classic II) coupled with acoustically matched TDH-39 supra aural headphone and Radio ear B-71 bone vibrator was utilized to estimate pure tone threshold, speech recognition threshold and speech identification score. A calibrated immittance meter (Amplaid-760) was used for obtaining tympanometry and acoustic reflex. A Desktop Computer of Core 2 Duo processor with adobe audition (version 3.0) software was used to record and present the developed test material.

Procedure:

After the estimation of pure tone threshold, speech reception threshold (SRT) and speech identification scores (SIS), the high frequency word identification list developed in Phase 1 was played (through computer plugged to audiometer) at 40 dB SL (wrf: SRT), delivering the stimulus through headphones. An external output of the audiometer was calibrated prior to testing each participant to 0 VU, using a 1000 Hz calibration tone. All participants were tested monaurally using the developed lists. The order of the lists was randomized to avoid order effect and an open set response in the form of oral response was obtained. All participant responses were scored by a native Mizo interpreter throughout data collection. Responses were only marked correct if they matched the target stimuli in both pronunciation and lexical tone. Prior to the administration of word recognition test, each participant was given the following instructions in English or in Mizo: " You will hear lists of Mizo words a number of times at a specific level. One word will be followed by another word. The silent gap between each word will be about 5 seconds. Please listen carefully and loudly repeat out the word that you hear. If you are unsure of a word, you are encouraged to guess. If you have no guess say, *I don't know*, or wait silently for the next word. Do you have any questions?"

Scoring:

After the estimation of MHF-SIT the responses were scored. Each correct response was given a score of 1. Each incorrect response was given a score of 0. The raw score was converted to percentage as below:

$$\text{Score}(\%) = \frac{\text{Total number of correct responses}}{\text{Total number of words presented}} \times 100$$

Statistical Analysis:

Statistical Package for the Social Sciences (version 16) software was used to carry out the statistical analysis. Descriptive statistics to find out the Mean and Standard Deviation, Independent *t*-test to compare the Mean scores of variables, one way ANOVA, repeated measure ANOVA and Bonferroni post hoc test were the statistical test used.

Results and Discussion

The results of the present study are described and reported under the following sub-headings.

Development of the High Frequency Word List/s

Selection of the words and their familiarity:

Initially a total of 445 words were collected which included both monosyllabic and bisyllabic words. These words were then rated for their familiarity using 10 native (5 female & 5 male) speakers of Mizo. As a result, out of 445 words 378 were rated as most familiar. Only these 378 words were considered for the development of Mizo high frequency speech identification test (MHF-SIT).

Results of LTASS of Target words:

LTASS results showed that out of 378 words, 250 words had highest energy concentration at and above 1 kHz. These 250 words were considered for the development of the final list. The 250 words considered were further categorized based on the different cut off peak frequency (1.0, 1.5, 2.0, 2.5, & 3.0 kHz) as given in the Table 1.

Table 1: Cut off frequency of selected 250 words

Cut off Frequency(kHz)	Number of words
1.0	250
1.5	213
2.0	186
2.5	124
3.0	68

Table 2: Mean and Standard Deviation (SD) of spectral peak frequency on selected word list.

List	Number of words	Mean Peak frequency (Hz)	SD
List 1: Half list 1	25	2937.68	421.96
List 1: Half list 2	25	2920.44	374.92
List 2: Half list 1	25	2926.36	450.45
List 2:Half list 2	25	2928.72	568.13
List 3: Half list 1	25	2946.60	587.99
List 3: Half list 2	25	2948.64	546.70
Total	150	2934.74	489.83

From the data given in Table 1, it is clear that 250 words had predominant spectral information above 1 kHz. It is also evident that as the cut-off frequency increased, the number of words decreased.

Although all 250 words showed spectral information higher than 1 kHz, in the present study only those words that have cut off frequency at and above 2 kHz were included for the final construction of the list. Sinthiya (2009) statistically compared spectral peak frequency between regular standardized phonetically balanced speech identification test and High frequency Speech Identification Test in Tamil. Results indicated that words with predominant spectral information above 2 kHz were more sensitive in detecting speech perception deficits in individuals with high frequency hearing loss, compared to words with energy above 1 kHz. Hence, the decision to use words with cut-off frequency above 2 kHz as inclusion criteria for the construction of final word lists is justified. Thus, there were 186 words available with peak frequency above 2 kHz, of which 150 words were randomly selected for the final list. These 150 words consisted of 100 monosyllabic and 50 bisyllabic high frequency words.

Construction of the Word subtests:

Word subtest was constructed from selected 150 words consisted of 100 monosyllabic and 50 bisyllabic words. Thus, based on this, three separate lists were prepared; List 1 and List 2 had 50 monosyllabic words each while the List 3 had 50 bisyllabic words. This was done because the redundancy in bisyllabic words could be more than that of monosyllabic words which could affect speech identification scores (Hirsh, Silverman, Reynolds, Eldert & Benson, 1952). It was also presumed that depending on the level of word difficulty in these lists speech identification scores could differ and a normative developed combining these words within the same list may be erroneous. Hence, separate monosyllabic and bisyllabic list were prepared.

Further, 50 monosyllabic words in both List 1 and List 2 were divided into two half list with 25 words each, thus, there were 4 half list of 25 words each. Similarly, List 3was also divided into two half list with 25 words each. This was done to provide a shorter version of the test which could be useful when the complete list cannot be administered due to time constraints. While dividing the lists, attempt was made to keep the frequency of high frequency sounds similar in all the half lists.

Descriptive statistics was done to find out the Mean and Standard Deviation (SD) of the spectral peak of 150 words selected for the final list. Table2 shows the Mean and Standard Deviation of spectral peak frequency obtained from LTASS of all the 6 lists.

Later, one way ANOVA was done to compare the equality in peak frequency energy concentration across 6 lists. The results indicated that there was no significant difference between the spectral peak frequencies of the words in the lists [$F (5, 144) = 0.013$, $p>0.05$] indicating that all 6 lists had similar high frequency energy concentration.

Development of the Normative for the Mizo High Frequency Speech Identification Test (MHF-SIT)

Normative was established for the developed test on 100 normal hearing individuals (50 Females & 50 Males) who were native speakers of Mizo. Speech identification scores were obtained for all 6 half lists for each ear separately. Words were presented randomly to both ears. The Mean and standard deviation of speech identification scores obtained for six lists are given in the Table3.

Comparison between the Ears:

Speech identification scores obtained for the left and right ear was compared for all the six lists (Table 3). It was observed that the majority of the normal hearing individuals obtained almost 100% speech identification scores in the left as well as right ear. Later, one way ANOVA (Table 4) was performed for the purpose of checking the effect of list, ear and interaction between the two. Results of one way ANOVA showed no significant difference between the speech identification scores obtained in two ears. Hence, the data from left and right ear were combined for further statistical analysis.

The present result showed that normal hearing individuals obtain almost 100% speech identification score on

Table 3: Mean and Standard Deviation (SD) of MHF-SIT scores in normal hearing individuals

Lists	Ears	Mean (%)	Range	1 SD
List-1 Half list 1	Right	99.5	100 - 96	1.25
	Left	99.5	100 - 96	1.30
	Total	99.5	100 -96	1.27
List-1 Half list 2	Right	99.6	100 - 96	1.10
	Left	99.6	100 - 96	1.20
	Total	99.6	100 - 96	1.17
List-2 Half list 1	Right	99.7	100 - 96	1.09
	Left	99.6	100 - 96	1.15
	Total	99.6	100 - 96	1.11
List-2 Half list 2	Right	99.5	100 - 96	1.13
	Left	99.6	100 - 96	1.20
	Total	99.5	100 - 96	1.25
List-3 Half list 1	Right	99.8	100 - 96	0.685
	Left	99.9	100 - 96	0.562
	Total	99.9	100 - 98	0.626
List-3 Half list 2	Right	99.8	100 - 96	0.876
	Left	99.8	100 - 92	0.971
	Total	99.8	100 - 92	0.922

Table 4: One-way ANOVA results across lists

List	F	p
Word-List-1 Half list 1	0.049	0.826
Word-List-1 Half list 2	0.058	0.881
Word-List-2 Half list 1	0.064	0.801
Word-List-2 Half list 2	0.202	0.653
Word-List-3 Half list 1	0.203	0.653
Word-list-3 Half list 2	0.093	0.760

high frequency words. This finding is in agreement with several earlier studies (Schwartz & Surr, 1979; Mascarenhas, 2002; Sudipta, 2006; Sinthiya, 2009; Ratnakar, 2010). The lowest score obtained among the 100 participants was 92%. Thus, it can be inferred that the specificity of the Mizo High Frequency Speech Identification Test (MHF-SIT) is good. Earlier studies (Schwartz &Surr, 1979; Mascarenhas, 2002; Sudipta, 2006; Ratnakar, 2010) have checked for the sensitivity with high frequency sensorineural hearing loss. However, due to time constraints that was not among the objectives of the present study.

Comparison between Males and Females

The speech identification scores were also compared across the 6 lists for males and females (50 males & 50 females). One way ANOVA was performed to see whether there was a significant difference in the speech identification scores between male and female participants. To do this, speech identification scores of both ears (right and left ears) obtained from males were compared with that of females separately for each of the 6 lists. The results revealed that there was no significant difference between the speech identification scores obtained between males and females in all the 6 lists as depicted in Table 5. This finding was in agreement with earlier studies (Kruger, 2010) where comparison was

Table 5: One way ANOVA for Speech Identification Scores between male and female participants

	Gender	Ear	Mean (%)	SD	p
List-1: Halflist-1	Female	Right	99.6	1.21	0.75
		Left	99.3	1.48	0.22
	Male	Right	99.5	1.31	0.75
		Left	99.6	1.09	0.22
List-1: Halflist-2	Female	Right	99.6	1.09	0.73
		Left	99.7	0.95	0.18
	Male	Right	99.6	1.21	0.73
		Left	99.4	1.40	0.18
List-2: Halflist-1	Female	Right	99.5	1.31	0.14
		Left	99.5	1.31	0.29
	Male	Right	99.8	0.79	0.14
		Left	99.7	0.95	0.29
List-2: Halflist-2	Female	Right	99.4	1.40	0.54
		Left	99.4	1.40	0.18
	Male	Right	99.6	1.21	0.54
		Left	99.7	0.95	0.18
List-3: Halflist-1	Female	Right	99.8	0.79	0.56
		Left	99.8	0.79	0.15
	Male	*Right*	99.9	0.56	0.56
		Left	100	0.00	0.15
List-3: Halflist-2	Female	Right	99.7	0.95	0.65
		Left	99.8	0.79	1.00
	Male	Right	99.8	0.79	0.65
		Left	99.8	1.13	1.00

Table 6: Results of Boneferroni's post hoc test showing the pair wise comparison across the word subtests

Lists	List-1 Half list 1	List-1 Half list 2	List-2 Half list 1	List-2 Half list 2	List-3 Half list 1	List-3 Half list 2
List-1Half list 1	-	NS	NS	NS	S	NS
List-1Half list 2	NS	-	NS	NS	S	NS
List-2Half list 1	NS	NS	-	NS	NS	NS
List-2Half list 2	NS	NS	NS	-	S	NS
List-3Half list 1	S	S	NS	S	-	NS
List-3Half list 2	NS	NS	NS	NS	NS	-

done to determine to see the presence of any significant statistical differences between males and females on word recognition scores in native speakers of Samoan and results showed no significant difference.

Comparison of Speech Identification Scores across lists to check the equality of different lists

To verify whether there is any significant difference in the speech identification scores across the 6 lists, repeated measures ANOVA was done. The results of repeated measure ANOVA indicated that there was a significant main effect of word list on speech identification scores [F (5, 995) = 4.106, $p < 0.01$]. To obtain a pair wise comparison across the 6 lists, Boneferroni's post hoc analysis was carried out. The results of Boneferroni's test are shown in the Table 6.

Results showed no significant difference in speech identification scores between 4 monosyllabic half lists, and between 2 bisyllabic half lists. Also, there was no significant difference between monosyllabic word (wordlist-2) half list 1 and the bisyllabic word lists in terms of speech identification scores. However, significant difference was seen between bisyllabic word lists (Halflist-1 of List-3) [$p < 0.05$] and other word lists, with bisyllabic word lists showing better speech identification scores as compared to monosyllabic word lists. This improvement in speech identification scores for bisyllabic word list could be because of the redundancy of the bisyllabic words, which is easier to identify as compared to monosyllabic words. A similar result was also found in earlier studies done by Hirsh et al. (1952).

Thus, although the results revealed that there was statistical difference, inspection of the mean speech identification scores obtained for the monosyllabic word lists and the bisyllabic word lists was above 99%. This indicates that the magnitude of the mean difference is very small and will not have any clinical importance. Hence, it can be concluded that any of the 6 word subtests can be used to assess the high frequency speech identification.

Conclusions

The results showed that the normal hearing participants obtained a mean score of more than 99% for all the 6 lists. No significant difference across the ear and gender was observed in speech identification scores across the lists. Further, pair wise comparison across the 6 half lists showed that any of the lists are equally comparable and can be used to obtain high frequency speech identification scores. However, the utility of test needs to be checked on individuals with high frequency hearing loss.

The Mizo High frequency Speech Identification Test (MHF-SIT) will be useful to identify and evaluate speech perceptual deficits in individuals with high frequency hearing loss. Speech perception abilities obtain using this test will give a better estimate of the communication handicap that these individuals possess as compared to regular speech test. This could also be useful in the selection of appropriate amplification devices for individuals with high frequency hearing loss and auditory training of high frequency words.

References

Abrol, G. M. (1971), Establishment of a pilot rehabilitation unit in Audiology and Speech Pathology in India, *Final report, New Delhi,* AIIMS.

American Speech-Language and Hearing Association. (1988). Guidelines for determining threshold level for speech. *American Speech-Language Hearing Association, 30*(3), 85-89.

Borden, G. J., & Harris, K. S. (1980). *Speech Science Primer: Physiology, acoustics and perception of speech.* Baltimore, MD: Williams and Wilkins.

Carhart, R. (1952). Speech audiometry in clinical evaluation, *ActaOtolarynglogia, 41*, 18 - 42.

Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure tone thresholds. *Journal of Speech and Hearing Disorders, 24,* 33-345.

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

Chung, D. Y., & Mack, B. (1979). The effect of masking by noise on word discrimination scores in listeners with normal hearing and with noise-induced hearing loss. *Scandinavian Audiology, 8,* 139-143.

Cooper, F. S., Delattre, P. C., Libermann, A. M., Borst, J. M., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds, *Journal of Acoustical Society of America, 24,* 597 - 606.

Corso, J. (1959). Age and sex differences in pure-tone thresholds. *Journal of the Acoustical Society of America, 31,* 498-507.

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

Egan, J. P. (1979). Basic aspects of speech audiometry. *Ear, Nose, and Throat Journal, 58*(5), 190-193. Retrieved from http://www.entjournal.com/ME2/Default.asp

Epstein, A. (1978). Speech audiometry. *Otolaryngologic Clinics of North America, 11*(3), 667-676. Retrieved from http://www.oto.theclinics.com.

Gardner, H. J. (1971). Application of high frequency consonant discrimination word list in hearing aid evaluation, *Journal of Speech and Hearing Disorders, 36*(3), 354 - 355.

Gelfand, S. A. (1998). Optimizing the reliability of speech recognition scores. *Journal of Speech Language and Hearing Research, 41*(5), 1088-1102.

Greer, L. F. (1997). *Performance intensity functions for digitally recorded Italian speech audiometry materials*. Unpublished master's thesis. Brigham Young University, Provo, UT.

Harris, R. W., & Christensen, L. K. (1996). *Spanish speech audiometry materials* [*Compact Disc*]. Provo, UT: Brigham Young University.

Harris, R. W., Goffi, M. V. S., Pedalini, M. E. B., Merrill, A., & Gygi, M. A. (2001). Psychometrically equivalent Brazilian Portuguese bisyllabic word recognition materialsspoken by male and female talkers. *ProFono, 13*(2), 249-262. Retrieved fromhttp://www.profono.com.br/

Harris, R. W., Nielson, W. S., McPherson, D. L., & Skarzynski, H. (2004). Psychometrically equivalent Polish monosyllabic word recognition materials spoken by male and female talkers. *Audiofonologia, 25,* 16-31.

Harris, R. W., Nissen, S. L., Pola, M. G., McPherson, D. L., Tavartkiladze, G. A., & Eggett, D. L. (2007). Psychometrically equivalent Russian speech audiometry materials by male and female talkers. *International Journal of Audiology, 46*(1), 47-66.

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*(3), 321-337.

Hughes, G. W., & Halle, M. (1956). Spectral properties of fricative consonants. *Journal of Acoustical Society of America, 28,* 303 - 310.

Jerger, J., & Jerger, S. (1971). Diagnostic significance of PB word functions. *Archives of Otoloaryngology, 93,* 573-580.

Kiukaanniemi, H., & Maatta, T. (1980). Speech discrimination and hearing loss sloping to high frequencies. *Scandinavian Audiology, 9,* 235-242.

Kruger, L. K. (2010). *Samoan Speech Audiometry: Developing Word Recognition Materials for Native Speakers of Samoan.* Unpublished Master's Thesis Brigham Young University, Brigham.

Lacroix, P.G., & Harris, J.D. (1979). Effects of high-frequency cue reduction on the comprehension of distorted speech. *Journal of Speech and Hearing Disorders, 44,* 236-246.

Mascarenhas, K. (2002). High Frequency-Kannada Speech Identification Test. *Unpublished Master's Dissertation,* University of Mysore, Mysore.

Martin, M. (1987). *Speech Audiometry.* London. Whurr Publications.

Martin, F. N. (1994). *Introduction to Audiology.* , New Jersey, Prentice Hall, Eaglewood Cliffs.

McDermontt, H. J., & Dean, M. R., (2000). Speech perception with steeply sloping hearing loss: effects of frequency transposition. *British Journal of Audiology, 34,* 353-361.

Marshall, L., and Bacon, S. P. (1981). Prediction of speech discrimination scores from audiometric data. *Journal of Speech and Hearing Research, 2*, 148-155.

Mayadevi (1974). *Development and standardization of common speech discrimination test for Indians.* Unpublished Master's Dissertation University of Mysore, Mysore.

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

Owens, E., & Schubert, E.D. (1968). The development of consonant items for speech discrimination testing. *Journal of Speech and Hearing Research, 11,* 656-667.

Owens, E., & Schubert, E. D. (1977). Development of the California consonant test. *Journal of speech and Hearing Research, 20,* 463-474.

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

Pavlovic, C.V. (1987). Derivation of primary parameters and procedures for use in speech intelligibility predictions. *Journal of the Acoustical Society of America, 82,* 413 - 422.

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/

Ramachandra (2001). *High frequency speech identification test in Hindi.* Unpublished Master's dissertation. University of Bangalore, India.

Ratnakar, Y. V. (2010). *A high frequency speech identification test in Telugu.* Unpublished Dissertation, University of Mysore, India.

Rout, A. (1996). *Perception of monosyllabic words in Indian children.* Unpublished Master's Dissertation, University of Mysore, Mysore.

Schuknecht, H. F. (1964). Further observations of the pathology of presbycusis. *Archive of Otolaryngology, 80,* 369-382.

Schwartz, D. M., & Surr, R. K. (1979). Three experiments on the California Consonant Test. *Journal of Speech and Hearing Disorders, 25,* 55-60.

Seaver, L. (2008). *The development of word recognition materials for native speakers of Tongan*. Unpublished master's thesis). Brigham Young University, Provo, UT.

Sinthiya, K. (2009). *A high frequency speech identification test in Tamil.* Unpublished Master's Dissertation. University of Mysore, India.

Stark, R. E. (1979). In L.J. Bradford and W.G. Hardy (Eds.) *Hearing and Hearing Impairment* (pp. 229-249). New York: Grune and Stratton.

Studebaker, G. A., Pavlovic, C.V., & Sherbecoe, R. L. (1987). A frequency importance for continuous discourse, *Journal of the Acoustical Society of America, 81*(4), 1130-1138.

Studebaker, G. A., & Sherbecoe, R. L. (2002). Audibility-Index functions for the connected speech test. *Ear and Hearing, 23*(5), 385-398.

Sudipta, K., & Yathiraj, A. (2006). *A high frequency English speech identification Test (HF-ESIT).* Student research at AIISH, Mysore (Articles based on dissertation done at AIISH), Vol, 4.

Turrini, M., Cutugno, F., Maturi, P., Prosser, S., Leoni, F. A., & Arslan, E. (1993). Bisyllabic words for speech audiometry: A new Italian material. *ActaOtorhinolaryngologica Italica, 13*(1), 63-77. Retrieved from http://www.actaitalica.it/.

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

Wang, M. D., Reed, C. M., & Bilger, R. C. (1978). A comparison of the effects of filtering and sensorineural hearing loss on patterns of consonant confusions. *Journal of Speech and Hearing Research, 21,* 5-36.

Williott, J. F. (1991). *Aging and the auditory system.* Singular, San Diego: Academic Press.

Wilson, R. H., & Strouse, A. (1999). Psychometrically equivalent spondaic words spoken by a female speaker. *Journal of Speech, Language, and Hearing Research, 42*(6), 1336-1346.

An Evaluation of Acoustic and Perceptual Effects of Feedback Management in Hearing Aids

[1]Kruthika S. & [2]Manjula P.

Abstract

The present study aimed at evaluating the efficacy of the feedback reduction methods in the hearing aid, namely the phase cancellation algorithm and the use of acoustic modification (damper) in the ear mould using insertion gain measure and Speech Identification Scores (SIS). The data were collected from 60 ears of 30 children who had severe to profound hearing losses in both ears. The results indicated that there was an increase in the gain available as well as the output with the activation of feedback management and with the use of dampers. The behavioural measure, the speech identification scores, was higher with the activation of feedback management method and dampers compared to without them. The increase in gain can be attributed to the principle of working of phase cancellation method which reduces the feedback so that useful gain can be increased. As the dampers smoothen the frequency response at mid frequencies, a greater available gain is possible at mid- and high- frequencies. The increase in the available gain along with the efficient cancellation of feedback may be attributed to improved speech identification scores in the condition with feedback management activated and with the use of dampers. The findings of the study support the necessity of use of feedback management method like phase cancellation algorithm in digital hearing aids and dampers in analog hearing aids in children with severe to profound hearing loss.

Keywords: *Feedback Management, Phase Cancellation, Dampers, Insertion Gain, Speech Identification Scores.*

Introduction

Audible feedback is amongst the most prominent problems with hearing aids (Kochkin, 2002). In a hearing aid, the acoustic feedback or squeal occurs when the output of the hearing aid leaks out of the ear canal and enters the hearing aid microphone and is amplified again. The acoustic leakage is often attenuated by the ear mould coupled to the hearing aid. The conditions necessary for audible feedback oscillation are met when the degree of attenuation is small and/or when the gain of the hearing aid is high (Kuk, Ludvigsen & Kaulberg, 2002).

Generally, this feedback is associated with high gain hearing instruments. During such times, the annoyance, frustrations and embarrassment caused by the feedback may even outweigh an individual's otherwise perceived benefit from amplification. Acoustic feedback also can indirectly reduce the benefit from amplification. The hearing aid users may prefer to opt less-than-optimal gain to avoid the likelihood of feedback, or use the hearing aids for situations known to be 'feedback-free', or in extreme cases, simply stop using the hearing instruments (Chalupper, Powers & Steinbuss, 2011).

Thus, acoustic feedback is annoying and reduces the maximum usable gain of the hearing devices (Siqueira, Speece, Petsalis, Alwan, Soli & Gao, 1996). These peaks in the response of the hearing aid, which are often high-frequency in nature, may produce an uncomfortable sharpness in the hearing aid processed speech and affect speech recognition (Freed & Soli, 2006). Acoustic feedback phenomenon can thus deteriorate the performance of digital hearing aids working at high gains, causing instability and speech degradation (Leira, Bueno, Pita, & Zurera, 2008). This phenomenon even contributes to the 'hearing aid effect' , as potential users of amplification view acoustic feedback as a part of the negative stigma (Cox & Alexander, 2000).

The main challenges faced by the hearing aid users prone to feedback problems are mostly threefold. First, it can distort the sound signal across all the frequencies, causing a noticeable reduction in sound quality. Second, it can be so annoying to the user's environment that he/she is forced to turn the instrument down, thus loosing the crucial speech information when it is most needed. And finally, it can restrict the full use of the volume control, which in turn limits the person's ability to hear and understand speech.

Acoustic feedback is also associated, more often, with children having severe to profound degree of hearing loss. In young children the problem of feedback is exacerbated as the external ear is still growing (Westwood & Bamford, 1995). Consequently, after certain weeks or months, the initially good fitted ear mould can become loose and hence may increase the probability of occurrence of feedback (Flynn & Flynn, 2006). For the above reasons, consideration on feedback management in hearing aids is of extreme importance in paediatric population with severe to profound hearing impairment.

Feedback reduction algorithms in digital hearing aids may provide a solution for some of these problems. The acoustic feedback suppression in hearing aids can in-

[1]Email: kruthika.aiish@gmail.com
[2]Professor in Audiology, Email:manjulap21@hotmail.com

crease the maximum insertion gain of the aid. The ability to achieve target insertion gain leads to better utilization of the speech band-width. Thus, improved speech intelligibility for the hearing aid user can be expected as the most probable outcome (Siqueira et al., 1996).

Different approaches to feedback management have been introduced in hearing instrument technology (Dillon, 2001). With the advent of digital signal processing, audible feedback oscillation can be minimized without sacrificing gain, audibility, loudness, and speech intelligibility. A more promising solution for acoustic feedback would be the use of a feedback cancellation algorithm. The feedback canceller produces an estimate of the feedback signal and subtracts this estimate from the microphone signal, so that, ideally, only the desired signal is preserved at the input. Since the acoustic path between the loudspeaker and the microphone can vary significantly depending on the acoustical environment, the feedback management must be adaptive.

Among the feedback reduction methods, generally phase cancellation has been used. Phase cancellation systems are capable of suppressing feedback without degrading the audibility of speech, and therefore, this type of feedback reduction is preferable (Chalupper, et al., 2011).

Techniques for better ear mould design such as reducing the size of the conventional vents, coupling of hearing aids, and fitting methods have also been proposed to reduce feedback problems (Cox, 1982). Use of dampers in ear moulds can increase the usable gain of hearing aids. They increase the low frequency attenuation and provide greater high frequency output and a smoother frequency response by reducing the peakedness caused at high frequencies due to feedback (Valente, 1984). Therefore aim of the present study was to evaluate the feedback reduction methods in hearing aids, namely the feedback reduction algorithm and use of acoustic modification (damper). The specific objectives were, first, to evaluate the effect of the feedback reduction strategies in hearing aids, such as the phase cancellation method and use of damper in ear mould, on the insertion gain measures. Second, to evaluate the effect of the feedback reduction method and use of damper in the ear mould, on speech identification scores (SIS).

Method

The study aimed to evaluate the efficiency of phase cancellation method and use of dampers as the two feedback management methods using insertion gain measures and behavioural measures.

Participants

The data were collected from a total of 60 ears of 30 children, in the age range of six to eight years (Mean age of 6 years 5 months). The participants considered for the current study had Kannada as their mother tongue and had pre-lingually acquired bilateral severe to profound hearing loss, pure tone average ranging from 75 to 120 dB HL in the speech frequencies. All of them had flat or gradually sloping (with a slope of <15 dB per octave) hearing loss in both the ears. On immittance evaluation with GSI Tympstar middle ear analyzer (version 2), all the participants got 'A' type tympanogram with reflexes being absent. TEOAEs carried out through ILO 292 instrument, were absent in both the ears revealing outer hair cell dysfunction in both the ears. Auditory Brainstem Responses done with Intelligent Hearing System were absent in both ears for all the participants.

The thresholds for the frequencies at which the feedback occurred mostly from 1500 Hz to 6000 Hz (Martin, & Robert, 2006) was equal to or greater than 90 dBHL for all the participants, irrespective of minimal residual hearing at low frequencies till 1000 Hz. All the participants were using binaural digital Behind-The-Ear (BTE) hearing aids with a gain lesser than the target gain due to the occurrence of the feedback. With this gain setting, all the participants obtained an aided closed-set speech identification (through picture identification task) score of 50% or greater in the aided condition. The participants had no significant history of otologic, neurologic, cognitive or psychological problems. All the participants attended listening training and speech therapy for a period of at least of three months and they had the auditory skills at least for the identification of words.

All the testing was carried out in an air-conditioned sound treated double room situation. The study was carried out in three different phases. Phase I- Selection of participants, Phase II-Insertion Gain measurement and Phase III-Aided behavioural testing.

Phase I: Selection of Participants

Audiological evaluation: A detailed case history, routine audiological tests including pure tone audiometry, speech audiometry and immittance evaluation were carried out for all the participants for each test ear to confirm the inclusion criteria.

Speech Identification Scores (SIS) for selection of participants: A digitally programmable two channel Behind-The-Ear hearing aid with a fitting range for severe-to-profound sloping hearing loss, with custom made soft shell ear mould was used for the testing. The hearing aid used had 2 channels and 8 bands with 4 programmable memories. The hearing aid had a maximum output level of 135 dB SPL, maximum gain of 70 dB and a reference test gain of 52 dB. The basic frequency response was from 200 Hz to 6400 Hz. The hearing aid utilized 'Active Feedback Intercept' which worked on the principle of phase cancellation method for feedback management. (Engebreston & St. George, 1993).

The hearing aid was programmed to match the target gain curves obtained using the proprietary prescription formula. A personal computer with NOAH-3 and hearing aid specific softwares, and Hearing Instrument Programmer (Hi-Pro) interface were used to program the hearing aids and to activate or de-activate the feedback reduction algorithm. The aided closed-set speech identification scores were calculated as the number of words correctly identified, out of a total of the 25 words, when presented at 45 dBHL. The response mode was through the picture identification. A score of 50\% and above was considered as the criterion for the inclusion of participants in the current study.

Phase II: Insertion Gain Measurement Procedure

Fonix 7000 real ear measurement (computer controlled real-time analyzer version 1.70) with probe tube microphone option was used in order to measure the amount of gain/output delivered by the hearing aid in the ear canal of the participant through the insertion gain measurement procedure.

The loudspeaker of the real ear measurement system was placed at approximately 12 inches and at 450 Azimuth from the participant, at the level of participant's head. The integrated probe microphone was placed on the test ear of the participant. The reference microphone was secured on the ear hanger above the ear. After the probe tube was inserted, the probe microphone body was pivoted towards the ear to help hold the probe tube in place.

The ear mould was placed next to the probe tube, so that the tube rested along the bottom of the canal part of the ear mould, with the tube extending at least 5 mm from the tip of the ear canal opening, where a marking was done. For the aided testing, length of the canal portion of the custom ear mould, in addition to a length of 5 mm, was considered as the marking point on the probe tube which was made to coincide at the tragal notch of the participant's test ear.

Real Ear Measurement Procedure: The sound field was levelled by keeping the probe tube near the ear canal. A digi-speech signal was used at 65 dB SPL for the measurement of Real Ear Unaided Response (REUR). A hearing aid was fitted with custom ear moulds. The NAL-NL1 fitting formula was used to prescribe the hearing aid gain. The hearing aid was programmed to match the target gain. During this process, if there was occurrence of feedback, the volume control was reduced to a level at which there was no feedback with the programmed gain. Thus, the real ear measurements were obtained at the reduced volume control setting in the 'feedback reduction - off' condition. The Real Ear Aided Response (REAR) was measured at this setting of the hearing aid. Further, closed-set Speech Identification Score was also obtained with these settings on the hearing aid.

Likewise, the hearing aid was programmed to reach the target gain with the 'feedback reduction - on' condition. The volume control was set to the optimum setting i.e., to a point where there was feedback. The Real Ear Aided Response (REAR) was measured with this setting. It must be noted that the same volume control setting was used for measuring the SIS from the participant.

All the participants were tested with the 'Active Feedback Intercept - on' and in 'Active Feedback Intercept - off' conditions. Later, a damper of resistance 4700 Ω was inserted into the tubing of the ear mould connected to the hearing aid at a distance of 9 mm from the tubing end. The damper was placed as close as possible to the earhook end of the hearing aid. With the use of dampers, the volume control setting that was attainable without the occurrence of feedback was noted and REAR was obtained with this volume setting without the activation of feedback management method. It must be noted that the same volume control setting was used for measuring the SIS from the participant.

The Real Ear Insertion Gain (REIG) was obtained which is the difference between the REAR and REUR across all the frequencies for the three aided conditions, namely with and without feedback management and with damper, for each test ear. Two other measures were included which were calculated based on the real ear gain obtained at different frequencies. The two measures were the High Frequency Average Real Ear Insertion Gain (HFAREIG in dB) and Added Stable Gain (ASG in dB). The high frequency average real ear insertion gain (HFAREIG) was obtained by averaging the real ear gain values across the speech frequencies, i.e., 1000 Hz, 1600 Hz and 2500 Hz. It was calculated with the assumption that the high frequency average gave a better estimate of speech perception abilities than other frequency averages (Lenzen, 2008). The ASG in dB was calculated by subtracting the REAG obtained in 'without feedback management' condition from REAG obtained in 'with feedback management condition' i.e., REAG (WFBM) - REAG (WOFBM) or subtracting REAG in 'without feedback management' condition from REAG in 'with damper' condition i.e., REAG (WDAMP) - REAG (WOFBM). This measure was obtained as it could be used as a quantitative measure to compare the benefit from different feedback management options, in comparison with no feedback management. Also, the effect of the available gain on the improvement in speech identification scores could be quantified.

Phase III: Behavioural Testing

The closed-set Speech Identification Scores (SIS) were obtained in quiet through monitored live voice presentation of the phonemically balanced word lists for children developed by Vandana (1998). For this, the gain was set at just below the level causing feedback. The stimuli were presented through a loudspeaker of the audiometer from 00 Azimuth placed at a distance of 1 meter from the head of the participant at 45 dBHL. The response mode was pointing to the appropriate picture out of a group of four pictures. The scoring was done based on the number of words correctly identified out of the total number of 25 words presented. This was done for feedback suppression algorithm activated and subsequently with the use of dampers with appropriately adjusted volume control settings. Thus, for each participant, the data on REAR, in dB SPL; HFAREIG, in dB); Added Stable Gain (ASG, in dB and Aided SIS (maximum score being 25) were collected.

Statistical Analysis

Appropriate statistical analysis was carried out for the data to verify the objectives of the study. The mean and standard deviation of the REAR (in dB SPL), HFAREIG (in dB), ASG in dB and SIS (with maximum score of 25) were obtained. The scores obtained using REAR (in dB SPL), ASG, HFAREIG (in dB) were compared across the three aided conditions (WOFBM, WFBM & WDAMP) and were compared across eleven discrete frequencies from 200 Hz to 8000 Hz. Data on SIS were compared across the three aided conditions (WOFBM, WFBM and WDAMP) to check for the significant differences, if any.

Results and Discussion

The data collected were tabulated and subjected to statistical analysis using Statistical Package for Social Sciences (SPSS 17.0 for windows version). Descriptive statistics and analysis of variance were computed to evaluate the objectives of the study. The results are discussed under Insertion Gain Measures and Behavioural Measure.

Insertion Gain Measures

The data on insertion gain measure obtained for all the 60 ears were analyzed in the three aided conditions, viz., without feedback management with the feedback management and by inserting the damper in the ear mould.

Real Ear Aided Response, REAR (in dB SPL)

Real Ear Aided Response (in dB SPL) was obtained at eleven discrete frequencies across the frequency range from 200 Hz to 8000 Hz, for 60 ears in three aided conditions. Descriptive statistics was used to compare the mean and standard deviation measures of the REAR values (Table 1) in the three aided conditions. Figure 1 shows the mean REAR values across the eleven frequencies for the three conditions (WOFBM, WFBM and WDAMP).

From Tables 1 and Figure 1, it is evident that the mean REAR values were highest for WFBM condition followed by WDAMP condition and then by WOFBM condition, across all the frequencies except for 1000 Hz and 7000 Hz. Further, Table 2 shows the mean REAR values of different frequencies in the three aided conditions.

As indicated in Table 2, a high mean REAR value was evidenced in the condition with feedback management (WFBM) compared to without feedback management (WOFBM) and with damper (WDAMP) conditions. To determine if this difference in REAR in the three aided conditions was significant, one-way repeated measure ANOVA was done. Results revealed a high significant difference between the three aided conditions [$F(2,118) = 113.55$, $p<0.01$]. As a high significant difference was evident on the repeated measure ANOVA, Bonferroni's multiple group comparison was carried out to evaluate the pairs of conditions which showed a significant difference. Table 3 shows the results of Bonferroni multiple group comparison showing the level of significance for the three pairs of conditions.

As indicated in Table 3, there was a highly significant difference between the WOFBM and WFBM conditions, and WFBM and WDAMP ($p<0.01$). In addition, significant difference was also noted between WDAMP and WOFBM ($p<0.05$).

The observed differences between the conditions along with WFBM having a significantly greater output compared to other two conditions (WOFBM and WDAMP) may be attributed to a greater available gain; and hence a greater output with the activation of the feedback management (Freed & Soli, 2006). These differences in the REAR is possible due to the use of digital technology in hearing aids, because of which mathematical estimations of the feedback path can be made and used to compensate for the feedback, essentially without affecting the input signal, while ideally preserving the desired output. Such a method provides an added 6 to 10 dB average headroom improvement and possibly avail more useable gain compared to without the feedback management activated (Olson, Musch & Struck, 2001).

The type of feedback management method used in the hearing aid in the present study was 'Active Feedback Intercept' which is based on the principle of working of phase cancellation algorithms. Since phase cancellation algorithms simply cancel out unwanted feedback, there is no gain reduction associated with elimination of feedback. On the contrary, this technique results in ASG, i.e., an increase in maximum gain with feedback management enabled compared to that with feedback management disabled. Thus, there was an increase in the output with the activation of feedback management method (Freed & Soli, 2006; Merks, Banerjee & Trine, 2006). The mean REAR values were higher

Table 1: Mean and Standard Deviation (SD) values of REAR (in dB SPL) for the three aided conditions (WOFBM, WFBM and WDAMP), across eleven discrete frequencies from 200 Hz to 8000 Hz (N=60 ears)

Conditions	REAR in dB SPL - Mean(S.D)										
	200 Hz	500 Hz	1000 Hz	1500 Hz	2000 Hz	3000Hz	4000 Hz	5000 Hz	6000 Hz	7000 Hz	8000 Hz
WOFBM	87.42(5.76)	95.98(6.71)	102.99(6.23)	102.02(7.90)	103.54(6.73)	93.45(5.97)	87.21(7.22)	80.89(8.51)	73.15(9.07)	67.96(9.76)	56.79(7.96)
WFBM	95.55(5.69)	105.24(5.57)	110.14(6.23)	109.15(5.71)	111.99(5.72)	101.26(5.26)	96.61(5.89)	92.51(7.46)	84.20(7.43)	77.32(10.13)	68.44(9.88)
WDAMP	90.43(6.72)	96.99(6.01)	100.14(6.15)	106.24(6.44)	107.71(6.50)	94.88(6.20)	90.95(6.99)	83.41(8.41)	78.68(9.12)	67.59(10.56)	61.13(10.8)

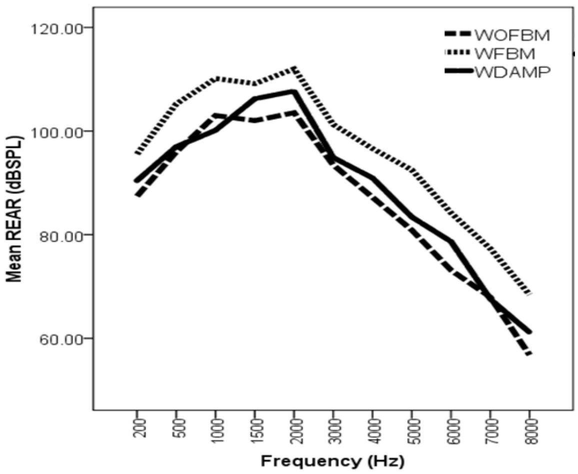


Figure 1: Mean REAR (in dB SPL) for the three aided conditions (WOFBM, WFBM and WDAMP).

Table 2: Results of descriptive statistics for Rear Ear Aided Response (REAR in dB SPL) indicating mean and SD values across the three aided conditions (WOFBM, WFBM and WDAMP)

REAR (in dB SPL)	
Conditions	Mean across frequencies(Standard Deviation)
Without feedback management (WOFBM)	86.49(15.42)
With feedback management (WFBM)	95.67(14.14)
With damper (WDAMP)	88.92(14.99)

Table 3: Results of Bonferroni multiple group comparison showing the level of significance across the pairs of conditions

Difference between conditions	p values
WOFBM & WFBM	0.000*
WFBM & WDAMP	0.000*
WDAMP & WOFBM	0.025**

*Note: *: $p < 0.01$=highly significant difference. **: $p < 0.05$= significant difference*

in the aided condition 'with dampers' for all the frequencies compared to 'without feedback management' method. This increase in the REAR values (and thus the real ear insertion gain values) may be because the dampers give a higher gain at higher frequencies and smoothen the frequency response at mid- to high- frequencies where the resonances caused by feedback results in sharp peaks in the frequency response (Valente, 1984). In the WOFBM method, the prescribed gain is reduced till the point where feedback does not occur, the gain across all the frequencies will be effectively lesser compared to WDAMP condition. Hence, a reduced output and thus the gain is the most likely outcome in the Aided condition without feedback management. The dampers reduce the sharp peaks caused due to the feedback and hence they reduce the gain and output SPL especially at low- to mid- frequencies (Valente, 1984). Hence, there are reduced REAR values in WDAMP condition compared to WFBM condition.

Also, frequency differences were noted for REAR (in dB SPL) across the three aided conditions. There was a greater output at mid- and high- frequencies from 500 Hz to 4000 Hz (with an average REAR of 105.73 dB SPL) compared to lower frequencies below 500 Hz (with an REAR of 95.55 dB SPL). Because the frequency response becomes sharper with the occurrence of feedback, there may be a significant difference of REAR across the frequencies in the without feedback management condition (Valente, 1984).

Because the feedback management methods provide added stable gain, the frequency response changes accordingly across the two conditions, viz., with and without feedback management conditions (Freed & Soli, 2006; Merks et al., 2006). Since the phase cancellation method is functional at 1500 Hz to 6000 Hz and more effectively it operates at 3000 Hz to 6000 Hz (Dyrlund, Henningsen, Bisgaard & Jensen, 1994), the peaks caused due to the presence of feedback are reduced. And this explains the observed differences in REAR values across the conditions for different frequencies. According to Valente (1984), there will be smoothening and hence a change in the frequency response with the use of dampers compared to WOFBM and WFBM, hence a significant difference is expected across the three conditions.

There are only certain frequencies where the feedback management method operates (Dyrlund et.al., 1994) and only certain frequency range that is smoothened by dampers (Valente, 1984). In addition, the frequency range where the feedback management method and dampers operate, may be different, that is there may be a difference in the REAR values across the frequency range for the three aided conditions.

A similar finding was noted in a study by Kuk and Ludvigsen (2002), who reported an increase in the available gain and hence the output across the frequency range of 200 Hz to 8000 Hz with the phase cancellation method. According to Martin and Robert (2006), phase cancellation not only preserves gain, but also because of its increased feedback margin, makes approximately 10 to 15 dB SPL more amplification available in the mid- to high- frequencies.

In a study by Lenzen (2008), the mean ASG ranged from 1.6 dB for low frequency band to 2.8 dB for the high frequency band. The mean difference of 1.2 dB in ASG between the low frequency band and the mid-frequency band was statistically significant (two tailed proportion $p<0.001$). Also, it was reported that the mean difference of 0.9 dB in ASG between the mid-frequency band and the high frequency band was statistically significant (two tailed proportion ($p<0.001$). There were no significant differences in mean ASG between the low frequency and high frequency band. The reason attributed to the reduced added stable gain values was that the maximum gain was reached initially and hence further improvement was not effective due to the 'ceiling effect'.

The ear mould dampers have an effect of smoothening the peaks from 1000 Hz to 3000 Hz (Taylor & Teter, 2009). As a result of this, the peaks are reduced and a smoother frequency response with higher gain is possible at mid frequencies (Dillon, 2001).

High Frequency Average Real Ear Insertion Gain (HFAREIG) calculated for the frequencies 1, 1.6 and 2.5 kHz, in all the three conditions

The data was obtained for all the three conditions across the frequencies 1, 1.6 and 2.5 kHz for 60 ears. Table 4 shows the descriptive statistics for the HFAREIG values at 1000 Hz, 1600 Hz and 2500 Hz obtained across the three aided conditions (WOFBM, WFBM and WDAMP).

Table 4 indicates that the mean HFAREIG value for WFBM condition is greater than the mean HFAREIG values for WOFBM and WDAMP conditions. Repeated measure ANOVA was done to find out the significant differences across conditions (WOFBM, WFBM and WDAMP), if any. A high significant difference was noted on repeated measure ANOVA for HFAREIG

Table 4: Descriptive statistics showing Mean and Standard Deviation for the HFAREIG values across the three conditions (WOFBM, WFBM and WDAMP)

HFAREIG (in dB)	
Conditions	Mean(Standard Deviation)
Without feedback management	34.99 (6.09)
With feedback management	42.04 (5.25)
With damper	37.69 (4.34)

Table 5: Results of Bonferroni's pair-wise comparison for HFAREIG across the three conditions (WOFBM, WFBM and WDAMP)

Conditions	Level of Significance HFAREIG (dB) (p)
WOFBM & WFBM	0.000*
WFBM & WDAMP	0.000*
WDAMP & WOFBM	0.002*

*Note: *: $p < 0.01$=high significant difference*

across the three conditions, WOFBM, WFBM and WDAMP [$F(2,118) = 55.92$, $p < 0.01$]. Bonferroni's pair-wise comparison was done to assess the significant differences between the HFAREIG for different conditions. Table 5 indicates results of Bonferroni's pair-wise comparison done across the three conditions for HFAREIG.

A highly significant difference was present across the three pairs of conditions for HFAREIG values as revealed from Table 5. The significant difference across the conditions may be attributed to the frequency range at which the feedback management is functional. It is supported by the fact that most of the feedback management is activated at frequencies between 1500 Hz to 6000 Hz. Accordingly, there would be a gain enhancement in this frequency range. This reason can be attributed to the observed differences between the HFAREIG values in WOFBM and WFBM conditions.

A study by Flynn and Flynn (2006) showed greater available gain with feedback management strategy at frequencies from 1.5 kHz to 3 kHz and thus might have resulted in a significant difference between the two conditions.

The present study shows an increase in HFAREIG by 7 dB, with and without feedback management conditions. This supports the findings of the study by Merks et al. (2006). He compared the feedback reduction performance of two hearing aids on 20 ears in terms of ASG derived by subtracting the Maximum Stable Gain (MSG) with feedback management deactivated from

Table 6: Mean ASG (in dB) across eleven discrete frequencies for WFBM and WDAMP conditions

Mean ASG for conditions(in dB)	Frequencies (in Hz)										
	200	500	1000	1500	2000	3000	4000	5000	6000	7000	8000
WFBM	8.13	9.26	7.14	8.44	7.80	9.40	11.61	11.04	11.04	9.36	11.64
WDAMP	3.01	1.01	-2.85	4.16	1.42	3.74	2.52	3.52	3.52	-0.37	4.33

that when it was activated. The MSG was calculated by averaging the real ear gain at 1 kHz, 1.6 kHz, and 2.5 kHz. The authors found that ASG ranged between 9 and 12 dB across the two hearing aids. The average difference in ASG between the two hearing aids used in the study was 3 dB (Lenzen, 2008).

The difference in HFAREIG values for WOFBM and WDAMP conditions was 2.7 dB. There was a high significant difference between WOFBM and WDAMP conditions which may be because dampers decrease the gain and the maximum output (Valente, 1984). Since they are more effective in reducing the peaks from 1 to 3 kHz, there will be reduction in the gain at this frequency range, compared to that obtained from WFBM condition. However, the gain reduction with the use of 'yellow' colour coded dampers was on an average 9 dB (Valente 1984), which was lesser compared to the gain reduction caused in an attempt to reduce the feedback in WOFBM condition.

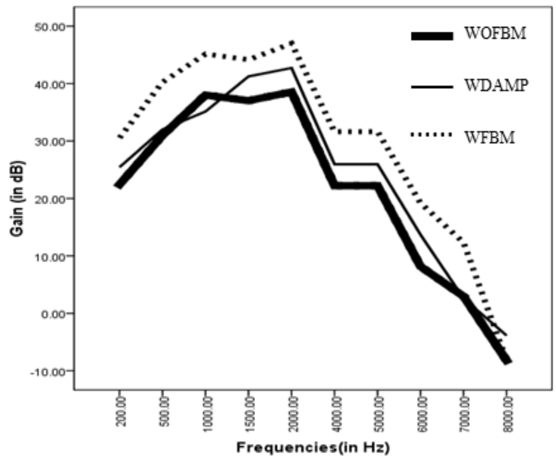


Figure 2: Gain (in dB) across the frequency range from 200 Hz to 8000 Hz, for the three conditions (WOFBM, WFBM and WDAMP)

Added Stable Gain (ASG) across the frequency range from 200 to 8000 Hz (for eleven discrete frequencies

As the ASG gives an idea of an increase in the available gain with feedback management and with dampers, comparison of the obtained ASG was made to account for the efficacy of the feedback management methods. Tables 6 shows the ASG for the two conditions (WFBM and WDAMP), across the eleven frequencies. Figure 2 shows the gain (in dB) across the frequencies for the three conditions (WOFBM, WFBM and WDAMP). Through this, the added stable gain can be calculated across the frequency range.

Figure 2 indicates that the gain for WFBM was greater compared to without feedback management and with damper. This finding was evident across the frequency range from 200 Hz to 8000 Hz. Significant differences between the two conditions were determined using pairwise comparison, if indicated.

Table 6 indicates that the mean ASG for WFBM condition was greater than WDAMP condition across all the frequencies. Also, mean ASG values were found to be greater for higher frequencies compared to mid- and low- frequencies. To find the average value of ASG across the frequencies, descriptive statistics was used. Table 7 shows the results of descriptive statistics giving mean and standard deviation for the ASG averaged across the eleven frequencies.

Table 7: Mean and SD (in brackets) values of Added Stable Gain (ASG) for the two conditions (WFBM and WDAMP

ASG (in dB)	
Conditions	Mean(Standard Deviation)
With feedback management	9.17(1.65)
With damper	2.24(2.26)

From Table 7, it is evident that the mean values for WFBM condition were greater than WDAMP. A difference of 6.0 to 7.0 dB was evidenced for ASG across the two conditions. Paired t-test was carried out to compare the ASG values for the two conditions WFBM and WDAMP. Results indicated a high significant difference between the two conditions on ASG values [t (10) =10.03, $p < 0.01$]. Several studies have reported the ASG values obtained with and without feedback management methods. The increase in ASG values with feedback management method can be attributed to the principle of working of the phase cancellation, which effectively reduces the feedback without reducing the gain. Moreover, it gives a greater available gain across the frequency range. Mean ASG values obtained through the method of feedback management varies across the frequencies. A few studies reporting the amount of ASG obtainable across the frequency range are discussed below.

According to Martin and Robert (2006), it was seen that the ASG was maximum at around 1 kHz and 3 kHz followed by higher frequencies. In addition, the study done by Martin and Robert (2006) reported greater ASG at higher frequencies (from 2 k to 5 kHz). Approximately 10 to 15 dB more amplification was made available in the mid-high frequencies through the phase cancellation method. According to Maxwell and Zurek (1995), the maximum added wideband stable gain was approximately 12 dB through the method of phase cancellation.

Field trials of a feedback-cancellation system built into a BTE hearing aid have shown increases of 8 to 10 dB in the gain used by individuals with severe hearing impairment (Bisgaard, 1993) and increases of 10 to 13 dB in the gain margin measured in real ears (Dyrlund et al., 1994). Computer simulations and prototype digital systems indicate that increases in gain between 6 and 20 dB can be achieved in an adaptive system before the onset of oscillation, and no loss of high-frequency response is noted (Engebretson & St.George, 1993).

Greenberg, Zurek, and Brantley (2000) reported that ASG ranged between -1 to 25 dB with a mean ASG of 8.5 dB for the experimental algorithm and approximately 5 dB for the other algorithms. Banerjee, Recker, and Paumen (2006) compared the feedback reduction performance of two hearing aids on 20 ears. Maximum stable gain (MSG) was calculated by averaging the gain at 1, 1.6, and 2.5 kHz. The authors found that the ASG ranged between 9 and 12 dB across the two hearing aids. However, a study by Lenzen (2008) indicated that the mean ASG ranged from 1.6 dB for low frequency band to 2.8 dB for the high frequency band. A mean difference of 1.2 dB in ASG between the low frequency band and the mid frequency band was statistically significant (two-tailed proportion $p < 0.001$). There were no significant differences in mean ASG between the low frequency and high frequency band.

Merks et al., (2006) did not report average ASG, but reported that ASG ranged from 3.5 to 16.3 dB. Banerjee et al. (2006) reported that ASG ranged between 2 to 18 dB with an average of 9 to 12 dB. Freed and Soli (2006) did not report average ASG, but reported that ASG ranged between 0 to 18 dB across all frequencies. Greenberg et al. (2000) reported an ASG ranging between -1 to 25 dB with an average ASG of 8.5 dB for one experimental algorithm and approximately 5 dB for the other experimental algorithms.

According to Kuk et al. (2002), there is an increase in Added Stable Gain in a wide range from 2 to 4 kHz, from as little as 8 dB to as much as 19 dB. An average of 12 to 13 dB was noted for the group. No increase in ASG was noted below 1 kHz, possibly because feedback being a high frequency phenomenon usually occurs above 1 kHz and target gain is typically reached below 1 kHz. Thus, there is probably no need for further increase in gain. Thus, the mean ASG values obtained in the present study was in accordance with the previous reports. However, the ASG obtained for the WDAMP condition was lesser compared with WFBM condition. This may be because of dampening effect from 1 to 3 k Hz which might have resulted in lower gain compared to the gain from WFBM condition.

Behavioural Measure

SIS values for the three conditions (WOFBM, WFBM and WDAMP)

The mean and standard deviation values of the Speech Identification Scores for each participant in three aided conditions were obtained. The maximum SIS was 25. Table 8 gives the mean and SD of the SIS.

Table 8: Mean and SD (SD) values of SIS (Max. 25) in three aided conditions (WOFBM, WFBM and WDAMP) WDAMP

	SIS
Conditions	Mean(Standard Deviation)
Without feedback management	17.45(2.15)
With feedback management	22.25(2.39)
With damper	20.78(2.54)

From Table 8, it is evident that the mean SIS for WFBM condition was greater than the SIS values for WOFBM and WDAMP conditions. It was found that the SIS was better in the WFBM condition than WOFBM and WDAMP conditions. Figure 3 shows the mean SIS scores with standard error (95% confidence level) across the three conditions.

Figure 3 shows the mean SIS values across the three aided conditions. The SIS in the WFBM condition was greater compared to SIS in the WDAMP and WOFBM conditions. In order to see if there was a significant difference, repeated measure ANOVA was done. It was found that there was a high significant difference between the three conditions on the SIS scores [$F(2, 118) = 275.85$, $p < 0.01$]. Following this, Bonferroni's pair-wise comparison was done to find the pair of conditions which showed a significant difference. Table 9 shows the results of Bonferroni's pair-wise comparison with significance levels.

Table 9 revealed that a highly significant difference existed across the conditions on SIS. This could be attributed to the added stable gain which was more in WOFBM condition than for the WFBM than WDAMP condition. The increase in ASG allows majority of

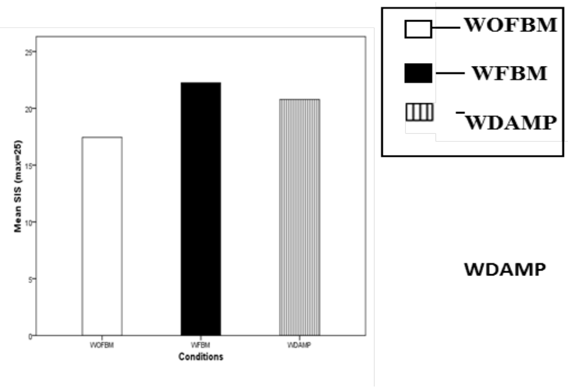


Figure 3: Mean SIS (Max.=25) for 60 ears across the three conditions (WOFBM, WFBM and WDAMP) (two-tailed with 95% confidence level).

wearers to achieve their desired gain without feedback in many more listening situations. The ability to use the target gain more consistently could result in better speech intelligibility, better sound quality, and a hassle-free listening experience.

Table 9: Results of Bonferroni's pair-wise comparison for SIS across three conditions (WOFBM, WFBM and WDAMP)

SIS scores across conditions	Level of Significance(p)
WOFBM & WFBM	0.000*
WFBM & WDAMP	0.000*
WDAMP & WOFBM	0.000*

*Note: *: $p < 0.01$=high significant difference*

In addition, the HFAREIG values were 7 dB greater for WFBM condition and 4.35 dB greater for WDAMP condition compared to WOFBM condition. The increase in the available gain at higher frequencies leads to better speech perception.

Dyrlund et al. (1994) reported that since there is a greater ASG at high frequencies, an improvement in speech identification performance is evidenced without causing feedback. Since feedback is a high frequency phenomenon, phase cancellation method would result in more of high frequency gain by cancelling the peaky responses (Martin & Robert, 2006). Christensen, Winfrey, and Stelmachowitz (2006) investigated the effectiveness of phase cancellation method and it was noted that a high frequency gain and improved perception of high frequency consonants resulted. The authors concluded that using feedback management helps to meet the mid- and high- frequency targets while providing maximum audibility for speech sounds. Hence, an improvement in SIS was noted in the WFBM condition.

The SIS in WDAMP condition was better than WOFBM condition. Peaks and troughs in the gain frequency response adversely affect the speech intelligibility and quality of amplified sound. Since smoothening of the frequency response takes place with the use of dampers and the peaks in the frequency response are reduced, an improvement in SIS was seen in WDAMP condition in comparison with WOFBM condition. However, the SIS obtained in WFBM condition was significantly better than the SIS obtained in WDAMP condition. This may be attributed to the added stable gain that is possible with the phase cancellation method. Where as in WDAMP condition, only the peaked responses are reduced providing a more stable output, without providing the ASG equivalent to that of the WFBM condition. This is expected since the purpose of the two is different. Hence, a significant difference was noted across the two conditions (WFBM and WDAMP).

Thus, the improvement on insertion gain measures with the activation of feedback management methods, viz., the feedback reduction strategy and the use of dampers, leads to a parallel improvement in terms of speech perception as well.

Conclusions

From the study it can be inferred that there is a highly significant improvement with the feedback management condition than without the feedback management in hearing aids. The improvement in the WFBM condition was due to the principle of working of phase cancellation method, because of which gain is not compromised while reducing the feedback. On the contrary, available gain increases with the activation of phase cancellation. Hence, greater output and gain value result with phase cancellation (Freed & Soli, 2006; Merks et al., 2006). With dampers, the amount of gain available and hence the output given also is significantly higher because of the effect of smoothening of the peaks in the frequency response. This leads to increased high- and mid- frequency response as the hearing aid wearers can increase the volume control. This is mainly because the peaks in the frequency response are reduced (Valente, Dunn & Roeser, 2000).

Increase in the Added Stable Gain and high frequency average real ear insertion gain led to better speech identification scores in feedback management activated condition than when it was de-activated. Dyrlund et al. (1994) reported that since there is a greater ASG at high frequencies, there is an improvement in speech identification performance without causing feedback. The SIS obtained in the WDAMP condition was better than in WOFBM condition, due to increased gain available in mid- and high- frequencies (Valente, Dunn & Roeser, 2000). However, the amount of gain available was lesser compared to WFBM condition.

References

American National Standards Institute. (1997). *Methods of measurement of real-ear performance characteristics of hearing aids.* ANSI S3.46-(1991). New York: American National Standards Institute.

Banerjee, S., Recker, K., & Paumen, A. (2006). A tale of two feedback cancellers. *The Hearing Review, 13*(7), 40 - 44.

Bisgaard, N. (1993). Digital feedback suppression - Clinical experiences with profoundly hearing impaired. In: Beilin J, Jensen GR. *Recent developments in hearing instrument technology.* 15th Danavox Symposium, Copenhagen, 370-384.

Chalupper, J., Powers, T. A., & Steinbuss, A. (2011).Combining phase cancellation, frequency shifting, and acoustic fingerprint for improved feedback suppression. *The Hearing Review,18*(1), 24-29.

Christensen, J. A., Winfrey, J. L., & Stelmachowicz, P. G. (2004). Applying adult hearing aid concepts to children: A feasibility study. *The Hearing Journal, 57*(4), 25-36.

Cox, R. M. (1982). Combined effects of earmold vents & suboscillatory feedback on hearing aid frequency response. *Ear and Hearing, 3*, 263-273.

Cox, R. M., & Alexander, G. C. (2000). Expectations about hearing aids and their relationship to fitting outcome. *Journal of the American Academy of Audiology, 11*(7), 368-382.

Dillon, H. (2001). *Hearing aids.* (2nd Ed.). New York NY: Thieme.

Dyrlund, O., Henningsen, L. B., Bisgaard, N., & Jensen, J. H. (1994). Digital Feedback Suppression (DFS): Characterization of feedback-margin improvements in a DFS hearing instrument. *Scandinavian Audiology, 23*(2), 135-138.

Engebreston, A. M., & St. George, F. M. (1993). Properties of an adaptive feedback equalization algorithm. *Journal of Rehabilitation Research Development, 30*(1), 8-16.

Flynn, M. C., & Flynn, T. C., (2006). Feedback cancellation, *The Hearing Journal, 59*(3), 58-63.

Freed, D. J., & Soli, S. D. (2006). An objective procedure for evaluation of adaptive antifeedback algorithms in hearing aids. *Ear and Hearing, 27*, 382-398.

Greenberg, J. E., Zurek, P. M., & Brantley, M. (2000). Evaluation of feedback-reduction algorithms for hearing aids. *Journal of the Acoustical Society of America*, *108*(5), 2366-2376.

Kochkin, S. (2002). MarkeTrak VI: 10-year customer satisfaction trends in the US hearing instrument market. *The Hearing Review, 9*(10), 14-25.

Kuk, F., Ludvigsen,C., & Kaulberg, T. (2002). Understanding feedback and digital feedback cancellation strategies. *The Hearing Review, 9*(2), 36-49.

Leira, M. A., Bueno, V. R., Pita, G. R., & Zurera, R. M. (2008). Acoustic feedback reduction based on FIR and IIR adaptive filters in ITE digital hearing aids. *Audio, Language and Image Processing,* 1442-1448.doi: 10.1109/ICALIP.2008.4590247.

Lenzen, N. M. (2008). *Differences in added stable gain between manufacturers, audiometric configurations, earmold styles, and frequency bands*. Unpublished Capstone Project submitted as a part of fulfillment for Doctor of Audiology to Washington University School of Medicine Program in Audiology and Communication Sciences, Washington.

Martin & Robert, L. (2006). Phase-cancellation feedback control: All hype aside, it's a big step forward. *The Hearing Journal, 59*(10), 56-58.

Maxwell, J. A., & Zurek, P. M. (1995). Reducing acoustic feedback in hearing aids. *IEEE Transactions on Speech and Audio Processing, 3*(4), 304-313. doi: 10.1109/89.397095

Merks, I., Banerjee, S., & Trine, T. (2006). Assessing the Effectiveness of Feedback cancellation in Hearing Aids. *The Hearing Review*, *13*(4), 53-57.

Olson, L., Musch, H., & Struck, C. (2001). Digital solutions for feedback control. *The Hearing Review, 8*(5), 44-49.

Siqueira, M. G., Speece, R., Petsalis, E., Alwan, A., Soli, S., & Gao, S. (1996). Subband adaptive filtering applied to acoustic feedback reduction in hearing aids. *Signals, Systems and Computers, 1*, 788-792.

Taylor, B., & Teter, D. (2009). Earmolds: practical considerations to improve performance in hearing aids. *Hearing review.* Retrieved from http://www.hearingreview.com /issues/articles/2009-09_01.asp

Valente, M. (1984). *Hearing aids.* (2nd Ed.). New York NY: Thieme.

Valente, M., Dunn, H. H., & Roeser, R. J. (2000). *Audiology Treatment.* New York NY: Thieme.

Vandana, S. (1998). Speech Identification Test For Kannada Speaking Children. Unpublished Independent Project submitted as a part of fulfillment of I M.Sc. (Speech & Hearing), to the University of Mysore, Mysore.

Westwood, G. F., & Bamford, J.M. (1995). Probe-tube microphone measures with very young infants: Real ear to coupler differences and longitudinal changes in real ear unaided response. *Ear and Hearing, 16*, 263-273.

Effect of Reverberation on Acceptable Noise Level in Individuals with Normal Hearing and Hearing Impairment

[1]Laxme Janardhanan & [2]N. Devi

Abstract

In the present study, the effect of changes in reverberation time on acceptable noise level in individuals with normal hearing and hearing impairment was investigated. The speech material for establishing acceptable noise level were digitally modified to create four reverberant conditions by applying different values of reverberation time (RT) to a non-reverberant condition (RT : 0, 0.4, 1.2, and 2 seconds). Two groups of 15 participants each (18-50 years) participated in the study; Group I had individuals with normal hearing, and Group II had individuals with bilateral mild- moderately severe hearing impairment, who had no previous experience with any amplification devices. Most comfortable listening level and background noise level measurements were established in each reverberant condition, and from these measurements, acceptable noise levels (ANLs) were calculated. In individuals with normal hearing, significant difference was found between ANL for non-reverberant stimuli and that for stimuli with RT of 2 seconds. This can be attributed to the unfavorable influence of reverberation on the primary talker, which might have interfered with the listeners' willingness to accept background noise. The ANLs of participants with normal hearing were better than aided and unaided ANLs of those with sensorineural hearing loss indicating the adverse effect of reverberation and noise on ANLs in participants with hearing impairment. The aided and unaided ANLs were significantly different, with a better value for aided ANL. This is clinically relevant as ANLs can be used to predict success with hearing aid.

Keywords:*Most comfortable listening level, background noise level, reverberant stimuli, sensorineural hearing loss.*

Introduction

Speech is seldom transmitted in a totally quiet, echo-free environment. Typically there are several degrading factors that can interact with each other to impede effective communication of speech. These factors include competing background noise, room reverberation, distance between speaker and listener, reduced hearing sensitivity, auditory processing abilities of the individual, etc.

Reverberation refers to the persistence or prolongation of sound within an enclosure as sound waves reflect off hard surfaces (Lochner & Burger, 1964; Nabelek & Pickett, 1974). Reverberation is present in all enclosed spaces to some degree (Lochner & Burger, 1964; Crandell & Smaldino, 2000). Reverberation is caused when reflections of sound waves of nearby surfaces create additional sound waves that overlap with the original signal. Reverberation time (RT) is defined as the time (seconds) it takes for the sound from a source to decrease in level by 60dB after the source has stopped (American National Standards Institute, 1970). A decrease of 60 dB represents a reduction of 1/1,000,000 of the original intensity of the sound. A longer RT results in more perceived reverberation, or echo on the part of the listener and has been well documented to produce a decrease in speech intelligibility (Houtgast & Steeneken, 1973; Duquesnoy & Plomp, 1980). Reverberation causes a prolongation of the spectral energy of the vowel sounds, which masks succeeding consonant phonemes, especially those consonants in word final positions. The masking effect of reverberation is more noticeable for vowels than for consonants because vowels exhibit greater overall power and are of longer duration than consonants. In effect, the distinct phonemes of speech become more difficult to discern, and the speech is therefore more difficult to understand (Houtgast & Steeneken, 1973). In highly reverberant environments, words may actually overlap with one another, thus causing reverberant sound energy to fill in temporal pauses between words and sentences.

Reverberation is more unfavorable when it occurs in combination with background noise than when present in isolation. Most listening situations have some noise and some degree of reverberation. When noise and reverberation are combined (as occurs frequently in actual listening situations) even younger listeners with normal auditory systems experience difficulty with speech understanding (Moncur & Dirks, 1967; Nabelek & Pickett, 1974).

The major sequelae of sensorineural hearing loss (SNHL) are speech perception difficulties, particularly in noisy or reverberant listening environments. Studies have noted marked variability on tasks of speech perception in reverberation, particularly among hearing-impaired (Nabelek & Pickett, 1974). As reverberation is not often confronted in the absence of background noise, it is important to evaluate amount of background noise that an individual would accept while listening to speech in reverberant conditions. Also, studies

[1]Email: laxme.aslp@gmail.com,
[2]Lecturer in Audiology, Email: deviaiish@gmail.com

have shown weak correlations between speech in noise scores, subjective evaluations of communication skills and hearing aid outcome. The acceptable noise level (ANL) measurement was first developed to quantify the amount of background noise an individual would accept while listening to continuous discourse (Nabelek, Tucker & Letowski, 1991).

According to Nabelek et al. (1991) to measure ANL, the listener's most comfortable listening level (MCL) for running discourse is measured. Next, the background noise level (BNL) is measured as the amount of background noise the individual is willing to accept while listening to the primary speech stimulus at MCL. Subtracting BNL from MCL gives ANL. The ANL measure assumes that speech understanding in the presence of noise may not be as important as the willingness to listen in the presence of noise.

Studies show that ANL is not related to hearing sensitivity (Nabelek, Tampas & Burchfield, 2004), gender (Rogers, Harkrider, Burchfield & Nabelek, 2003), and age (Nabelek, Freyaldenhoven, Tampas, Burchfield & Muenchen, 2006). ANL remains relatively constant when fitted with monaural or binaural amplification (Nabelek et al., 1991). Nabelek et al. (1991) found that ANL did not vary with different types of background noise. They used multi-talker babble, speech-spectrum noise, traffic noise, noise of a pneumatic drill, and music as background noise. Of the background noises used, music was the only one which showed significant effect. This is due to the variability of the music sample, the frequency spectrum of the music sample, and/or the listener's preference for the music sample. Plyler, Madix, Thelin and Johnston (2007) investigated the influence of high frequency information (i.e., beyond 2000Hz) on ANLs in individuals with normal hearing and impaired hearing. They reported that information beyond 2000 Hz may change (i.e., improve or degrade) some listeners' acceptance of background noise.

Studies which have been conducted to examine the physiological correlates of ANL suggest that ANL may be mediated by non-peripheral factors; it may be mediated, in part, beyond the level of the superior olivary complex where binaural processing initially occurs within the central auditory nervous system (Harkrider & Smith, 2005). Nabelek et al. (1991) speculated that ANL may be inherent to the individual.

Adams, Gordon-Hickey, Moore and Morlas (2010) evaluated the effects of reverberation on ANL in younger and older adults with normal hearing sensitivity. The results revealed no significant effect for age and/or reverberation time on MCL or ANL findings.

However, previous researches have not specifically examined how the presence of reverberation changes a hearing impaired individual's preferred listening level for speech, and the acceptance of noise. This is an important consideration as most listening environments are degraded by both noise and reverberation and the detrimental effect of combination of noise and reverberation affects individuals with hearing impaired more adversely than those with normal hearing. Therefore, the aim of the present study was to investigate the effect of reverberation and changes in reverberation time on ANL in individuals with normal hearing and hearing impairment. The study compared the effect of changes in reverberation time on ANL in individuals with hearing impairment under unaided and aided conditions. Unaided and aided ANL of individuals with hearing impairment were separately compared with the ANL of individuals with normal hearing.

Method

Participants

Thirty participants in the age range of 18 to 50 years were included in the study. They were divided into two groups; first group included fifteen individuals with normal hearing and second group included fifteen individuals with bilateral mild to moderately severe sensorineural hearing loss who had no previous experience with any amplification devices. All were native speakers of Kannada language. All the participants had speech identification scores of $> 75\%$ in both ears and negative history of middle ear infections, active speech and language disorder, and neurologic disorder or any cognitive listening deficits. They did not have any illness on the day of testing. Prior permission was taken from all the participants for their willingness to participate in the study.

Speech Material

The speech material used for determining ANL procedure were three standardized passages in Kannada compiled by Savithri and Jayaram (2005) and one passage developed by Sairam and Manjula (2002) which were spoken with normal vocal effort by a native female speaker of Kannada. Using this material, four experimental conditions were created, with varying amounts of reverberation. The unaltered condition had no reverberation (NR), while the remaining three conditions had varying amounts of digitally added reverberation created by applying RTs of 0.4 second (RT1), 1.2 seconds (RT2) and 2 seconds (RT3) to the NR stimuli which was accomplished by using Adobe Audition 1.5 multi-track sound editing software. The parameters like, the attack time and the high frequency absorption times were kept constant for all the three conditions of reverberations. These speech materials (primary talker) were recorded onto a Sony compact disc and were played through a personal computer, whose output was routed through the auxiliary input of the double channel audiometer. The speech material was presented through one channel of the audiometer at 0 azimuth.

Kannada speech babble developed by Anitha (2003) which was recorded onto a Sony compact disc was used as the background competing stimulus in the study. This was played through a personal computer, whose output was routed through the auxiliary input of the double channel audiometer. The speech babble was presented through the other channel of the audiometer at 180 azimuth.

Test Procedure

Preliminary procedures included otoscopy, and a behavioral audiometric evaluation. The audiometric testing was performed using a double channel audiometer calibrated according to ANSI S3.6 (1996) standards. Speech recognition thresholds and speech identification scores were also obtained. All testing was carried out in a double-room sound treated environment, with ambient noise level in the permissible limits as per ANSI S3.1 (1999).

To determine ANL, the conventional procedure (Nabelek et al., 1991) that involved the tester adjusting the level of the test words to the most comfortable listening level (MCL) of the participant was employed. The measures of MCL were obtained using a three part bracketing procedure. The primary discourse was presented at 30 dB HL and increased in 5 dB steps until the participant indicated that the speech was louder than the participant would want to listen to it. The primary discourse was then reduced in 5 dB steps until the participant indicated that the speech was "too soft." From this level, the level of the story was adjusted until the subject found his most comfortable listening level or the level he would want to "listen to the story on the radio." The primary discourse was then adjusted up and down in 2 dB steps until accurate MCL was established. After establishing MCLs, the BNLs were determined. The passages were played at the level of the MCL of the subject through loudspeakers at 0 azimuth and simultaneously, multitalker babble was presented through loudspeakers at 180 azimuth. The presentation level of the multitalker babble was 30dB HL and its level was increased by 5dB steps until a point at which the participant was willing to accept the noise without becoming tired or tensed while listening to and following the words of the passage. The noise was next adjusted up and down in 2 dB steps until the participant indicated that the 12-talker babble was at the highest level that was acceptable while listening to the story without becoming tired or tense. The maximum level at which he or she could accept the noise without becoming tired or tensed was considered as the BNL. The ANL (dB) was calculated as the difference between MCL (dBHL) and BNL (dBHL) for each participant. ANLs were calculated for individuals with normal hearing (ANL1) as well as for those with hearing impairment in unaided (ANL2) and aided (ANL3) conditions for non reverberated as well as for different reverberated stimuli. The aided ANLs were established after fitting the participants of the second group with a 5 channel digital behind the ear hearing aid which had a fitting range of mild to severe degree of hearing loss. Presentation of the reverberant condition was randomized for each participant. Additionally, rest periods was given to the participants as needed. The randomization and rest periods served to reduce the likelihood of practice or fatigue effects.

Results and Discussion

Effect of Reverberation on ANL in Participants with Normal Hearing (Group I)

Descriptive statistics was done to calculate the mean and standard deviation (SD) of MCL, BNL and ANL of individuals with normal hearing under different reverberant conditions. Table 1 shows the mean and standard deviation (SD) of MCL, BNL and ANL of participants

Table 1: Mean and Standard Deviation (SD) of MCL, BNL and ANL of Participants of Group I under Different Reverberant Conditions

	RT (seconds)	Conditions	Mean (dB)	SD
Group I (N=15)	NR	MCL	43.13	05.32
		BNL	36.20	05.80
		ANL1	06.93	03.45
	RT1	MCL	43.73	04.61
		BNL	37.13	04.75
		ANL1a	06.60	04.10
	RT2	MCL	45.27	04.67
		BNL	36.00	05.10
		ANL1b	09.27	05.12
	RT3	MCL	46.33	04.82
		BNL	33.53	05.77
		ANL1c	12.80	05.17

Table 2: Mean and Standard Deviation (SD) Of MCL, BNL and ANL of Participants of Group II under Non - Reverberant and Different Reverberant Conditions

	RT (seconds)	Conditions	Parameter	Mean(dB)	SD
Group II (N=15)	NR	Unaided	MCL	54.27	06.49
			BNL	42.47	06.79
			ANL	11.80	06.11
		Aided	MCL	43.13	04.31
			BNL	32.40	05.18
			ANL	10.73	04.88
	RT1	Unaided	MCL	54.07	06.04
			BNL	41.20	06.25
			ANL	12.87	04.96
		Aided	MCL	42.80	03.91
			BNL	34.20	03.99
			ANL	08.60	04.42
	RT2	Unaided	MCL	57.53	06.01
			BNL	38.60	06.23
			ANL	18.93	05.44
		Aided	MCL	47.13	04.26
			BNL	31.27	05.70
			ANL	15.87	05.66
	RT3	Unaided	MCL	63.47	05.63
			BNL	35.47	05.96
			ANL	28.00	04.05
		Aided	MCL	50.67	04.32
			BNL	29.73	05.16
			ANL	20.93	05.47

of Group I under different reverberant conditions.

In the present study, the ANLs varied from 3 to 14 dB with a mean (SD) of 6.93 dB (3.45) for non - reverberant stimulus. From Table 1, it can be observed that the mean ANL for non reverberant condition (ANL1) in individuals with normal hearing was 6.93 dB. The mean ANL for RT of 0.4 seconds (ANL1a), 1.2 seconds (ANL1b) and 2 seconds (ANL1c) were 6.60 dB, 9.27 dB and 12.80 dB respectively.

Repeated measure ANOVA was done to find out within condition effects of participants in Group I. Significant difference was noticed with $F(3, 42) = 32.066$, $p < 0.01$. Multivariate ANOVA was done to find out between condition effects of participants in Group I. Significant difference was noticed with $F(3, 12) = 37.635$, $p < 0.01$. In both the cases, Post hoc analysis was administered and the result showed significant difference between the conditions at $p < 0.01$ level.

Significant difference was found between ANL for non-reverberant stimulus and that for stimulus with RT of 2 seconds. The ANLs for stimuli with RT of 0.4, 1.2 and 2 seconds were significantly different from each other. But, ANLs for non- reverberant stimulus was not significantly different from ANL for stimulus with RT of 0.4 seconds. Even though the mean ANL for the stimulus with RT of 1.2 seconds was greater than that for the non-reverberant stimulus, there was no significant difference between the two conditions.

Gordon-Hickey and Moore (2008) reported that small, but significant changes in ANL occurred with a reduction in intelligibility of the primary discourse. The primary talker conditions in their study included the Arizona Travelogue with forward presentation (intelligible), reversed presentation (unintelligible), and Chinese discourse (unintelligible to study participants). The unintelligible conditions resulted in an increase in ANL of 1.5 to 2.2 dB. This finding is in contrary to other ANL studies, which have found that ANL is not related to scores of speech understanding in noise (Nabelek et al., 2004, 2006). Different studies have been reported that spectral smearing resulting from reverberation can adversely affect the intelligibility of speech (Houtgast & Steeneken, 1973; Nabelek, Letowski & Tucker, 1989). RTs approaching 1 second can negatively impact the intelligibility of the speech signal for individuals of all hearing abilities, particularly when background noise is present (Nabelek & Pickett, 1974; Sato, Sato, Morimoto & Ota, 2007). So, the RTs of 1.2 and 2 seconds included in the present study were large enough to impact the intelligibility of the speech signal.

One of the effects of reverberation on an acoustic signal is an overall increase in signal intensity (Houtgast & Steeneken, 1973; Finitzo-Hieber & Tillman, 1978), and hence, a decrease in MCL was hypothesized in reverberant conditions in comparison with MCL in the non-reverberant condition. It was also hypothesized that this anticipated change in MCL would result in a change in ANL. But, Franklin, Thelin, Nabelek and Burchfield (2006) had reported that with every 1 dB increase in the presentation level of the primary discourse, the ANL was decreased only by 0.25 dB. This suggested that with 2dB increase in intensity across the reverberant conditions, a maximum decrease in ANL would be only about 0.5 dB.

The increase in ANL with reverberation in the present study can be attributed to the unfavorable influence of reverberation on the primary talker, which might have interfered with the listeners' willingness to accept background noise. The 0.4 seconds reverberation is not functional enough to impede the primary talker and hence the difference between ANLs for non reverberant stimuli and 0.4 seconds reverberant stimuli is less compared to ANLs for non reverberant stimuli and stimuli with RT of 2 seconds. The absence of significant difference between ANL for RT of 1.2 seconds and that for non reverberant stimuli can be correlated to the BNL that the participants were able to put-up-with. The RT of 1.2 might not have interfered with the primary talker and hence the ANL for RT of 1.2 seconds is not significantly different from that for a non- reverberant primary talker.

Effect of Reverberation on ANL in Participants with Hearing Impairment (Group II)

Descriptive statistics was done to calculate the mean and standard deviation (SD) of MCL, BNL and ANL of individuals with hearing impairment under different reverberant conditions. Table 2 shows the mean and SD of MCL, BNL& ANL of participants of Group II under non reverberant and different reverberant conditions.

Unaided Condition: The unaided ANL varied from 5 to 23 dB with a mean (SD) of 11.80 (6.11) for non - reverberant condition. From Table 2, it can be noticed that in unaided condition, the mean ANL for non reverberant stimulus (ANL2) in individuals with hearing impairment was 11.80 dB. The unaided ANL for the stimulus with RT of 0.4 seconds (ANL2a), 1.2 seconds (ANL2b) and 2 seconds (ANL2c) were 12.87 dB, 18.93 dB and 28.00 dB respectively.

Repeated measure ANOVA was done to find out the within conditions effects of Group II (unaided). Significant difference was noticed with $F (3, 42) = 68.242$, $p < 0.01$. Multivariate ANOVA was done to find out the between conditions effects of Group II (unaided). Significant difference was noticed with $F (3, 12) = 55.038$, $p< 0.01$. In both the cases, Post hoc analysis (Bonferroni) was administered and the result showed significant difference between the conditions at $p< 0.01$ level.

Table 3: Comparison between ANLs of Participants of Group I and unaided ANLs of Participants of Group II

Group comparison	F value	Sig.
ANL1 vs. ANL2	7.215	.012
ANL1a vs. ANL2a	14.235	.001
ANL1b vs. ANL2b	25.098	.000
ANL1c vs. ANL2c	80.275	.000

Significant difference was present between ANLs for non-reverberant stimulus and those for stimulus with RT of 1.2 and 2 seconds. The ANLs for stimulus with RT of 0.4, 1.2 and 2 seconds were significantly different from each other. But ANLs for non- reverberant stimulus were not significantly different from those for stimulus with RT of 0.4 seconds. RTs approaching 1 second can negatively impact the intelligibility of the speech signal for individuals of all hearing abilities, particularly when background noise is present (Nabelek & Pickett, 1974; Sato et al., 2007). So, the RTs of 1.2 and 2 seconds included in the present study were large enough to impact the intelligibility of the speech signal.

Comparison of Unaided ANLs of Participants with Hearing Impairment with the ANLs of Those with Normal Hearing across Different Reverberant Conditions

One-way ANOVA was done to compare the unaided ANLs of participants with hearing impairment with the ANLs of those with normal hearing across different reverberant conditions. Table 3 shows the comparison of ANLs of participants of Group I with unaided ANLs of participants of Group II under non reverberant and different reverberant conditions. From the table, it can be

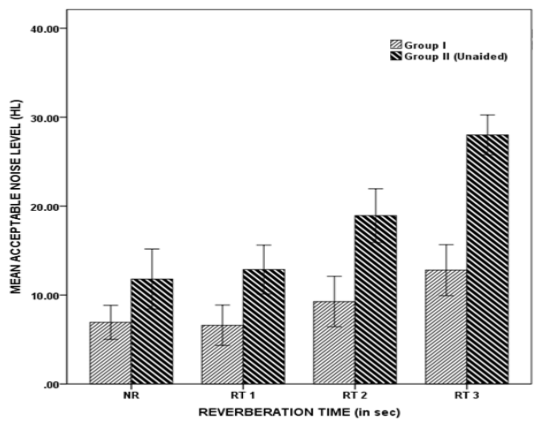


Figure 1: Mean ANL across Different RTs for Group I and Group II (Unaided).

Table 4: Comparison between ANLs of Participants of Group I and aided ANLs of Participants of Group II

Group comparison	F value	Sig.
ANL1 vs. ANL3	6.066	.020
ANL1a vs. ANL3a	1.650	.210
ANL1b vs. ANL3b	11.229	.002
ANL1c vs. ANL3c	17.511	.000

Table 5: Comparison between unaided and aided ANLs of Participants of Group II

Groups	t value	Sig.
ANL2 vs. ANL3	2.306	.037
ANL2a vs. ANL3a	3.043	.009
ANL2b vs. ANL3b	2.283	.039
ANL2c vs. ANL3c	4.899	.000

concluded that ANLs of participants of Group I was significantly different from the unaided ANLs of participants of Group II. Figure 1 represents the mean ANLs of Group I and Group II (unaided) across non- reverberant and different reverberant conditions. As per the results, there was significant difference between the ANLs of Group I and unaided ANLs of Group II in all reverberant conditions. This can be attributed to the fact that individuals with hearing impairment form a highly heterogeneous group, with great variability in speech understanding, particularly in degraded listening conditions such as noise and reverberation (Nabelek & Pickett, 1974; Nabelek & Letowski, 1985). The unaided ANL values are higher compared to those for participants in Group I with maximum difference at RT of 2 seconds, the condition in which the effect of reverberation is more adverse compared to other two conditions of reverberation.

Aided Condition: For the non - reverberated condition, the aided ANL varied from 5 dB to 20 dB with a mean (SD) of 10.73 (04.88). From Table 2, it can be noticed that in aided condition, the mean ANL for non reverberant stimulus (ANL3) was 10.73 dB. The aided ANL for RT of 0.4 seconds (ANL 3a), 1.2 seconds (ANL3b) and 2 seconds (ANL3c) were 8.60dB, 15.87 dB 20.93 dB respectively.

Repeated measure ANOVA was done to find out the within condition effects of Group II (aided). Significant difference was noticed with F (3, 42) = 36.885, $p < 0.01$. Also, Multivariate ANOVA was done to find out the between condition effects of Group II (aided). Significant difference was noticed with F (3, 12) = 23.957, $p < 0.01$. In both the cases, Post hoc analysis (Bonferroni) was administered and the result showed significant difference between the conditions at $p < 0.01$ level.

Significant difference was found between ANL for non-reverent stimulus and those for stimulus with RT of 1.2 and 2 seconds. The ANLs for stimulus with RT of 0.4, 1.2 and 2 seconds were significantly different from each other. But ANLs for non- reverberant stimulus was not significantly different from ANL for stimuli with RT of 0.4 seconds. This can also be attributed to the adverse influence of reverberation in the presence of background noise which might have reduced the listener's willingness to accept noise while following the words of the primary talker.

Comparison of Aided ANLs of Participants with Hearing Impairment with ANLs of Those with Normal Hearing across Different Reverberant Conditions

One way ANOVA was done to compare the aided ANLs of participants with hearing impairment with the ANLs of those with normal hearing across different reverberant conditions. Table 4 shows the comparison of aided ANLs of participants of Group II with ANLs of Group I participants under non- reverberant and different reverberant conditions. From Table 4, it can be concluded that ANLs of participants of Group I was significantly different from the aided ANLs of participants of Group II.

Figure 2 represents the mean ANLs of Group I and Group II (Aided) across non - reverberant and different reverberant conditions. As per the results, there was significant difference between the ANL of Group I and aided ANLs of Group II in all reverberant conditions except for RT of 0.4 seconds. That is, even with amplification, the performance of Group II participants did not approach the performance of Group I participants. The difference in ANL between Group I and Group II (aided) can be attributed to the lack of acclimatization of the participants of Group II with the hearing aid. Also, the hearing aids were programmed to match the target gain, along with considering the listening needs of

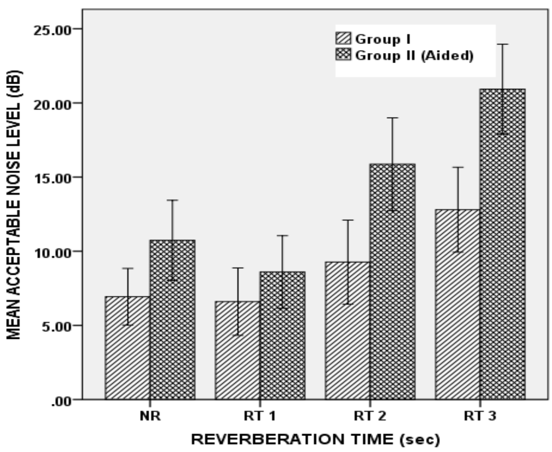


Figure 2: Mean ANL across Different RTs for Group I and Group II (Unaided).

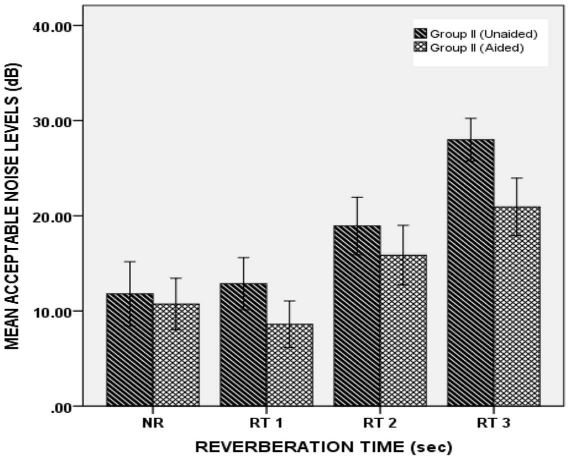


Figure 3: Mean ANL across Different RTs for Group I and Group II (Unaided).

the participants. No other special features for enhancing the listening in presence of reverberation or noise were activated in the hearing aid. This also might have contributed to the difference in ANLs across both the groups.

Comparison of Unaided and Aided ANLs of Participants with Hearing Impairment across Different Reverberant Conditions

Paired sample t test was administered to compare the aided and unaided ANLs of participants with hearing impairment across different reverberant conditions. Table 5 shows the comparison between unaided and aided ANLs of participants of Group II. From Table 5, it can be concluded that the unaided and aided ANLs of participants of Group II was significantly different from each other. Figure 3 represents the mean ANLs of Group II (Unaided & Aided) across non-reverberant and different reverberant conditions. As per the results, there was significant difference observed between the unaided and aided ANLs in all reverberant conditions. The mean difference was maximum for RT of 2 seconds. The aided performance was better even though it is not approaching the performance of participants with normal hearing. This suggests that the willingness to accept noise improved once the amplification was provided. Even though ANL is inherent to the individual (Nabelek et al., 1991), when the amplification is provided, the overall audibility is improved, and that might have contributed to the improvement in aided ANL values compared to unaided condition. But since reverberation is present, the unfavorable influence of reverberation on the willingness to accept background noise affects the individuals with SNHL more adversely than those with normal hearing. Therefore, even though the audibility is improved, the aided performance of individuals with hearing impairment in the presence of reverberation might not approach the ANL values of individuals with normal hearing, even though it is better than the unaided ANLs.

The ANL of all the groups varied between a wide range, which is in support with previous studies (Rogers et al., 2003 & Nabelek et al., 2006). In general, individuals with hearing impairment show great variability in speech understanding, particularly in degraded listening conditions such as noise and reverberation. The adverse effects of reverberation would have interfered with the primary talker used in the present study which would have lead to the increase in MCL as the reverberation time increased. When the primary talker is reverberated, the listener is expected to accept less level of the background noise so as to follow the words of the story at his comfortable level. So, this might have lead to the increase in MCL and decrease in BNL in conditions with reverberation. The 0.4 seconds reverberation is not functional enough to impede the primary talker and hence in cases of Groups I and II, the difference between ANLs for non - reverberant stimuli and 0.4 seconds reverberant stimuli is less compared to ANLs for non reverberant stimuli and stimuli with RT of 1.2 and 2 seconds.

The ANL of all the groups varied between a wide range, which is in support with previous studies (Rogers et al., 2003 & Nabelek et al., 2006). In general, individuals with hearing impairment show great variability in speech understanding, particularly in degraded listening conditions such as noise and reverberation. The adverse effects of reverberation would have interfered with the primary talker used in the present study which would have lead to the increase in MCL as the reverberation time increased. When the primary talker is reverberated, the listener is expected to accept less level of the background noise so as to follow the words of the story at his comfortable level. So, this might have lead to the increase in MCL and decrease in BNL in conditions with reverberation. The 0.4 seconds reverberation is not functional enough to impede the primary talker and hence in cases of normals and hearing impaired, the difference between ANLs for non - reverberant stimuli and 0.4 seconds reverberant stimuli is less compared to ANLs for non reverberant stimuli and stimuli with RT of 1.2 and 2 seconds.

Conclusions

The acceptance of background noise is reliant on the individual person and can unfailingly be tested without hearing aids (i.e., even before aids are recommended and fitted). Thus, the ANL may be beneficial as an indicator of eventual successful hearing-aid use. The ANL measures can be included in the routine audiological test battery as it gives idea regarding the success of hearing aid use and also, it does not take much time to administer (2-3 minutes). ANL values with reverberant primary talker have important application in the field of rehabilitation. The finding that reverberation has negative effect on individual's willingness to accept noise should be considered while designing the classrooms or therapy rooms. This study also emphasize that while programming a hearing aid 'echo stop' feature can be activated which can be used in reverberant environments, so that the willingness to accept noise can be improved.

As people with low ANL values will be successful hearing-aid use, finding an effective strategy to reduce listeners ANLs would increase their chances of benefiting from aural habilitation. Effect of reverberation on ANL needs to be further explored across different degrees of hearing loss, so that optimized and effective fitting can be achieved for individuals with varying degrees of hearing loss. Effect of reverberation on ANL needs to be further probed into in children with hearing sensitivity within normal limits and those with various types and degrees of hearing loss; ANL being an indicator of successful hearing aid usage, it needs to be practiced on a regular basis during the clinical fitting trials.

References

Adams, E. M., Gordon-Hickey, S., Moore, R. E., & Morlas, H. (2010). Effects of reverberation on acceptable noise level measurements in younger and older adults. *International Journal of Audiology* , *49*, 832-838.

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

American National Standards Institute (1999). *Maximum permissible ambient noise level for audiometric rooms.* (ANSI-S3.1-1999). New York: ANSI.

American National Standards Institute (1970). *Acoustical Terminology* (ANSI-S1.1-1970). New York: ANSI.

Anitha R. (2003). *The effect of speech babble of different languages on speech identification scores*. Unpublished independent project submitted to University of Mysore, Mysore.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools*, *31*, 362-370.

Duquesnoy, A. J., & Plomp, R. (1980). Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis. *Journal of the Acoustical Society of America, 68*(2), 537 - 44.

Franklin, C. A., Thelin, J. W., Nabelek, A. K., & Burchfield, S. B. (2006). The effect of speech presentation level on acceptance of background noise in listeners with normal hearing. *Journal of the American Academy of Audiology*, *17*(2), 141 - 146.

Gordon-Hickey, S., & Moore, R. E. (2008). Acceptance of noise with intelligible, reversed, and unfamiliar primary discourse. *American Journal of Audiology*, *17,* 129- 35.

Harkrider, A. W., & Smith, S. B. (2005). Acceptable noise level, phoneme recognition in noise, and measures of auditory efferent activity. *Journal of the American Academy of Audiology, 16*, 530-545.

Houtgast, T., & Steeneken, H. J. M. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acta Acoustica* , *28*, 66 - 73.

Lochner, J., & Burger, J. (1964). The influence of reflections in auditorium acoustics. *Journal of Sound and Vibration, 4,* 426 - 454.

Moncur, J., & Dirks, D. (1967). Binaural and monaural speech intelligibility in reverberation. *Journal of Speech and Hearing Research, 10*, 186-195.

Nabelek, A. K., & Pickett, J. M. (1974). Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. *Journal of the Acoustical Society of America, 56*(2)*,* 628 - 639.

Nabelek, A. K., & Letowski, T. R. (1985). Vowel confusions of hearing-impaired listeners under reverberant and nonreverberant conditions. *Journal of Speech and Hearing Disorders*, *50*, 126 -131.

Nabelek, A. K., Letowski, T. R., & Tucker, F. M. (1989). Reverberant overlap and self- masking in consonant identification. *Journal of the Acoustical Society of America, 86, 1259-65.*

Nabelek, A. K., Tucker, F. M., & Letowski, T.R. (1991). Toleration of background noises: Relationship with patterns of hearing aid use by elderly persons. *Journal of Speech and Hearing Research, 34*, 679 - 85.

Nabelek, A. K., Tampas, J.W., & Burchfield S.B. (2004). Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions. *Journal of Speech Language and Hearing Research, 47*, 1001 - 1011.

Nabelek, A. K., Freyaldenhoven, M. C., Tampas, J. W., Burchfield, S. B., & Muenchen, R. A. (2006). Acceptable noise level as a predictor of hearing aid use. *Journal of the American Academy of Audiology, 17,* 635-649.

Plyler, P. N., Madix, S. G., Thelin, J. W., & Johnston, K. W. (2007). Contribution of high-frequency information to the acceptance of background noise in listeners with normal and impaired hearing. *American Journal of Audiology, 16,* 149-156.

Rogers, D. S., Harkrider, A. W., Burchfield, S. B., & Nabelek, A. K. (2003). The influence of listener's gender on the acceptance of background noise. *Journal of American Academy of Audiology, 14* (7), 372 - 82.

Sairam (2002). Long Term Average Spectrum in Kannada. Independent project submitted to University of Mysore, Mysore.

Sato H., Sato H., Morimoto, M. & Ota, R. (2007). Acceptable range of speech level for both young and aged listeners in reverberant and quiet sound fields. *Journal of Acoustical Society of America, 122*(3), 1616 - 23.

Savithri, S. R., & Jayaram, M. (2005). Rate of Speech/ Reading in Dravidian Languages, AIISH Research Fund Project.

Effect of Music Exposure on Online Subcortical Plasticity

[1]Mahima Gupta & [2]Sandeep Maruthy

Abstract

The present study aimed to investigate influence of music exposure on the online plasticity of auditory system. The online plasticity was documented using context-dependent changes of speech evoked brainstem responses, which in turn was compared across musicians, music listeners and, non-music listeners. Speech perception in noise was recorded as a behavioral index of online plasticity. The experimental data was collected on 30 normal hearing adults. The results showed that speech perception in noise was better in musicians than that in other two groups. Whereas online plasticity was similar in the three groups. The enhanced speech perception in noise in musicians has been attributed to the training related changes in the olivocohlear bundle. The music exposure however, did not influence the online plasticity. The findings of the study support that only the active task like singing or playing an instrument is advantageous for speech perception in noise.

Keywords: *Online plasticity, speech ABR, speech perception in noise, plasticity*

Introduction

Animal experiments and human behavioral and electrophysiological studies have shown that the auditory cortex shows changes in plasticity, i.e. it is capable of reorganization as a function of experience (Tremblay, Kraus, Carrell & McGee, 1997). The term 'neural plasticity' refers to the alterations in the physiological and anatomical properties of neurons in the brain in association with sensory stimulation or deprivation. Studies have shown that both long-term and short-term experience affects the functioning of the brain (Shinn-Cunningham, 2001; Tremblay et al., 1997; Russo, Nicol, Zecker, Hayes & Kraus, 2005; Wong, Skoe, Russo, Dees & Kraus, 2007; Madhok & Maruthy, 2010). Long-term plasticity refers to the reorganization of the physiological and anatomical properties of brain neurons secondary to the training done for several months or years. Similarly, the changes that are resultant of few hours or days of training are referred as short-term plasticity.

In the past, plasticity was believed to be a phenomenon observed only in cortical structures, while the later experiments have evidenced plasticity even in the subcortical structures (Krishnan, Xu, Gandour & Cariani, 2005; Musacchia, Sams, Skoe & Kraus, 2007; Russo et al., 2005; Madhok & Maruthy, 2010).

Researches by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009) and, Skoe and Kraus (2010) reported the presence of a new type of plasticity which is termed as online plasticity. According to their findings, repetitive presentation of the stimulus induces online plasticity within few hours which causes the automatic sharpening of brainstem representation of speech cues related to voice pitch. This repetition induced neural fine tuning is found to be strongly associated with perception of speech in noise, suggesting that this type of plasticity is indeed functional (Chandrasekaran et al., 2009).

Skoe and Kraus (2010) demonstrated that human subcortical activity evolves in response to repetition of entire melody and repetition of a note within the melody within the ongoing stimulus stream. They found a robust enhancement to the repeated note appearing to develop monotonically over the 1.5 hour session. It was proposed by the authors that the subcortical online plasticity results from the statistical enhancement of intrinsic circuitry interacting with top-down influences such as auditory memory, musical knowledge, expectation and/or grouping via the corticofugal pathway. Hanan and Maruthy (2011) observed the presence of online plasticity only for spectrally dissimilar contextual stimulus and not for spectrally similar context.

The speech elicited ABR represents the pitch encoding (F0 & its harmonics) at the level of brainstem (Wong et al., 2007; Musacchia et al., 2007). It is shown that the encoding of pitch associated with complex sounds is due to the role of the neural phase-locked activity related to F0 (Swaminathan, Krishnan, Gandour & Xu, 2008). Bidelman, Gandour and Krishnan (2010) showed that the auditory brainstem encodes pitch irrespective of the context. Their findings also suggest that the pitch encoding is better in musicians than non-musicians for the linguistically and musically relevant features. Wong et al. (2007) found that the pitch encoding is better in musician group. They found correlation of the effect of long-term music training on linguistic pitch encoding, at the brainstem level. Hence, it could be said that the experience leads to superior representation of the pitch in native speakers and musicians. Review of literature (Musacchia et al., 2007; Wong et al., 2007) reveals that the speech ABRs are enhanced in musicians when compared to non-musicians. Furthermore, it is reported by the researchers (Parbery-Clark Skoe & Kraus, 2009; Musacchia et al., 2007; Wong et al., 2007) that there is direct relationship between the number of years of music training and the robustness of brainstem

[1]Email: gt.mahima@gmail.com,
[2]Reader in Audiology, Email: msandeepa@gmail.com

responses obtained, with the response being better with more years of practice.

In the real listening situations, it is very common to come across noisy situation. Thus, it becomes essential to understand the involvement of brainstem in the perception of speech in the presence of noise. Kumar and Vanaja (2004) suggested that the efferent auditory pathway plays an important role in the perception of speech in the presence of noise. Parbery-Clark et al. (2009) reported that if the brainstem responses evoked for speech in the presence of noise have early latencies, the HINT scores would also be good.

It is reported that the individuals with poor performance on HINT showed delayed latencies and lower magnitude for the formant transition in the presence of noise (Anderson, Skoe, Chandrasekaran & Kraus, 2010; Anderson, Skoe, Chandrasekaran, Zecker & Kraus, 2010). The poor temporal resolution at the brainstem is attributed to be the cause of the behavioral findings. Thus, speech ABR can be an electrophysiological index of deficits in speech perception in noise. The correlation between speech ABR and SPIN is an evidence for the role of brainstem in the SPIN.

Musicians are found to have enhanced spectral and temporal representation of the stimulus at the subcortical level. This enhancement is attributed to the presence of active top-down mechanisms, such as attention, memory, and context (Kraus, Skoe, Parebery-Clark & Ashley, 2009). Thus, based on the knowledge about these mechanisms and research support, it could be assumed that the auditory system, apart from showing long-term and short term plasticity, also shows plasticity for the stimulus presented for a very short duration known as online plasticity. Earlier, many studies have documented the longterm and short-term neuroplastic changes in musicians. However, there is a dearth of literature on the *online plasticity* in musicians. Considering that their corticofugal pathway in trained for duration of their music training, one would expect that the online plasticity is better in musicians. Furthermore, it is not clear in the previous studies whether the training related changes seen in musicians is due to formal practice of vocal or instrumental music which is an active task or, due to listening to music on a regular basis which is relatively a passive task. Therefore, to answer these questions, the present study was taken up. There were two specific objectives of the study that is to compare the online plasticity among musicians, music listeners and control individuals on an electrophysiological paradigm and, to compare the relationship between online plasticity and speech perception in noise.

Method

The present study was initiated with null hypotheses that there is no significant difference between the online plasticity among musicians, music listeners and non-music listeners. The following method was used to test the null hypothesis.

Participants

A total of 30 normal hearing (peripheral & central) adults, in the age range of 18 to 30 years participated in the present study. They were divided into three groups based on their music experience. Each group consisted of 10 participants. Group I consisted of participants who never received any formal music training and did not have the habit of listening to music on a regular basis. This group served as a control group. Group II had participants who had never received any formal music training, but would listen to music (either vocal or instrumental) at least for one hour a day and 5 days a week since last 5 years, at least. Group III had participants who had received formal music training for minimum of 5 years. The participants in this group had received either instrumental or vocal music training, and had been practicing the same everyday at least for an hour. These participants would also listen to music at least for an hour each day.

A checklist developed to profile their audiological status, and exposure in music, was administered prior to the actual testing. The interested readers can refer to the complete dissertation submitted to University of Mysore. Based on the responses obtained using this checklist, the groups were subdivided. A written consent was taken from all the participants before carrying out any of the tests.

Test Stimulus

The experimental procedure required presentation of two types of stimulus: context stimuli and a core stimulus. The contextual stimuli occurred more frequently than the core stimulus. The core stimulus was operationally termed so, as only the responses obtained for this particular stimulus was of interest in data analysis. The three stimuli used were, synthetically generated /da/, f2 filtered /da/, and the white noise. The /da/ stimulus was originally synthesized in Auditory Neuroscience lab, Northwestern University, Chicago by Professor Nina Kraus, Principal investigator, Auditory Neuroscience lab, Northwestern University, Chicago. The same stimulus was used in the present study with the consent of Prof. Kraus. The white noise of 40 ms duration was generated at Psychoacoustic Lab, All India Institute of Speech and Hearing, Mysore, using Adobe Audition (version 1.5) software.

All the three stimuli were individually normalized and then group normalized to obtain equal average RMS power of 93.4 dBSPL, using MATLAB software. They were then loaded into the personal computer with Bio-Logic Navigator Pro AEP Software (Version 7.0). The

Table 1: Mean behavioral thresholds in (dBSPL) of the synthetically generated syllables /da/, F2 filtered /da/ and, the white noise

Stimulus	Mean behavioral thresh-old(dBSPL)	Approximated mean behavioral thresh-olds(dBSPL)
Synthetically generated /da/	27.71	30
F2 filtered /da/	28.82	30
White noise	31.14	30

synthetic speech syllables /da/, and the filtered /da/ were subjected to a subjective rating of quality judgment from 15 sophisticated listeners with normal hearing. This was done for the nHL calibration. To do so, all the three stimuli were presented at a repetition rate of 10.9/s through the insert receivers of the Bio-Logic Navigator Pro AEP system. The mean behavioral thresholds obtained were as given in Table 1.

All tests were administered in acoustically treated rooms with noise levels at permissible limits (ANSI S3.1, 1991).

Test Procedure

Only the participants who fulfilled the inclusion criteria were subjected to the actual test procedure. The preliminary testing included pure tone audiometry, immittance evaluation, speech perception in noise (SPIN) test and auditory brainstem responses (ABR). The actual experimental paradigm involved recording of Speech perception in noise and speech evoked brainstem responses.

In the experimental testing, SPIN and speech evoked ABR were recorded. SPIN was assessed using the standardized English sentence speech stimuli developed by Thakur and Kumar (2008). Multi talker babble was used as the background noise, during the test administration. Target sentences were presented at 40dBHL. Speech to noise ratio (SNR) required to understand 50% of the words in sentences (SNR-50), was estimated. Level of the multi-talker babble was varied in 2dB steps using adaptive staircase procedure to yield 50% correct response. SNR was made adverse when the subject repeated all the key words in a sentence. Target sentences and noise were presented monaurally in the right ear only.

Speech evoked ABR was recorded in three different stimulus conditions. The target responses were recorded using vertex (non- inverting) and nape (inverting) while the baseline activity was recorded with ground on the left ear mastoid. The experiment involved presentation of stimuli in two stimulus paradigm that is, repetitive and odd ball paradigm. The stimulus used in the repetitive paradigm was synthetically generated /da/ stimulus, for which two recordings of 1500 sweeps each were done. This was done for establishing baseline for responses obtained in the contextual condition. In the odd ball paradigm, a core (infrequent) stimulus was presented in the presence of a contextual (frequent) stimulus. The synthetically generated /da/ was the core stimulus and was presented with the probability of 25%, against 75% for the frequent stimuli which was either white noise (in condition 2) or F2 filtered /da/ (in condition 3).

Response Analysis

The resultant averaged waveform had both transient and sustained components in it. The responses were analyzed subjectively as well as objectively. The transient responses were analyzed subjectively by two experienced audiologists to mark peak V, A, and C. The peak latency and amplitude were noted down at marked points. Marking of the peaks in a representative averaged waveform is shown in Figure 1.

Additionally, objective analysis was done for evaluating the spectral composition of sustained portions of the response using Fast Fourier transform (FFT). This was done using the MATLAB R 2009a platform and software (Brainstem toolbox) developed by Kraus (2004) at Northwestern University. Fourier analysis was performed on the 11.4 to 40.6 ms epoch of the FFR in order to assess the amount of activity occurring over three frequency ranges; (103-121Hz), (454-719Hz) and (721- 1155Hz). These frequency ranges were chosen because the neural responses at these frequencies correspond to the Fundamental frequency, first formant and higher harmonics of the stimulus /da/ respectively (Johnson et al., 2008). A 2 ms on 2 ms off Hanning ramp was applied to the waveform (to avoid the spectral splatter). Zero-padding was employed to increase the number of frequency points where spectral estimates were obtained. The raw amplitude value of the F0, F1 and higher frequency (HF) component of the FFR were then measured and noted. The FFR response in a representative spectrum is shown in Figure 2.

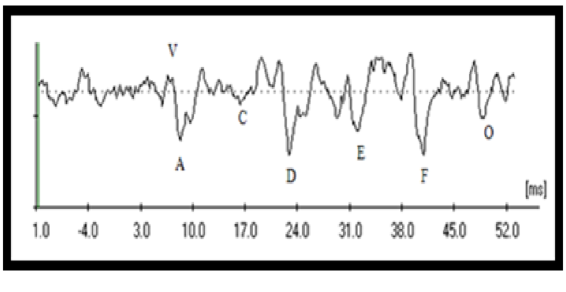


Figure 1: Representation of the marking of the peaks in a representative averaged waveform.

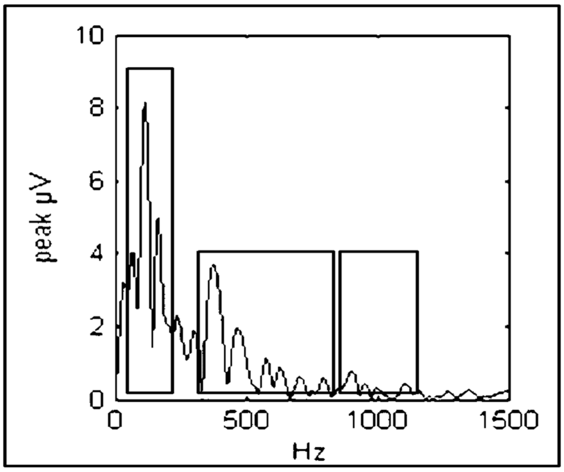


Figure 2: Representation of FFR responses in a representative spectrum.

Results

The mean latency and amplitudes derived were statistically compared to test the effect of condition and, group. The individual data of brainstem responses were also correlated with the respective speech perception in noise performance to understand the relationship between the two variables. All the statistical tests were performed using Statistical Package for Social Science software (version 16.0).

The percentage of occurrences of V, A, and C peaks among the participants of the three groups, in the three stimulus conditions is given in Table 2.

As evident from Table 2, the wave V and A were present in 100% of the participants in all three stimulus conditions, whereas this was not the casewith wave C. The absence of the peak was noticed for different conditions in different individuals, which reduced the actual number of data in Analysis of Variance (ANOVA). It was due to this reason that peak C was not included for the comparisons among the various conditions and groups.

Table 2: The percentage of occurrence of wave V, A and C among the participants of three groups, in the three stimulus conditions.

Group	Condition	Transient waves (in %)		
		V	A	C
NML	1	100	100	100
	2	100	100	100
	3	100	100	90
ML	1	100	100	80
	2	100	100	80
	3	100	100	80
MC	1	100	100	100
	2	100	100	100
	3	100	100	100

Table 3: The mean and standard deviation (SD) of peak latency (ms) of wave V, A, and C in the three stimulus conditions, for the three groups

Wave	Group	Condition		
		1	2	3
		Mean in ms (SD)	Mean in ms (SD)	Mean in ms (SD)
V	NML	7.49 (0.62)	7.79 (0.60)	7.78 (0.59)
	ML	7.36 (0.45)	7.66 (0.44)	7.69 (0.50)
	MC	7.23 (0.30)	7.39 (0.30)	7.33 (0.25)
A	NML	8.47 (0.64)	8.78 (0.55)	8.95 (0.58)
	ML	8.37 (0.54)	8.61 (0.50)	8.49 (0.56)
	MC	8.25 (0.37)	8.42 (0.36)	8.55 (0.38)
C	NML	17.51 (0.87)	17.81 (0.89)	17.73 (0.65)
	ML	17.64 (1.12)	17.72 (1.42)	17.91 (1.20)
	MC	16.96 (1.55)	17.27 (1.74)	17.04 (2.19)

Effect of Condition on Transient Responses

The mean and standard deviations of latency of transient responses were estimated among the three stimulus conditions (one repetitive paradigm & two oddball paradigms), using descriptive statistics. The data is as given in Table 3.

The data in Table 3, shows that the mean latencies were prolonged in condition 2 and condition 3 compared to that in condition 1. This is true for wave V, A, and C. To verify whether the observed mean differences in wave V and A are significantly different across the three stimulus conditions, repeated measure ANOVA was done taking group as between-subject variable. The results (Table 4) showed significant main effect of stimulus condition on the latency of wave V and, A. There was no significant interaction between condition and group.

Because there was significant main effect of stimulus condition on wave V and A, pair-wise comparison was

Table 4: The results of repeated measure ANOVA for wave V, and A latencies

Wave	Effect of condition		Condition X Group	
	F	df(error)	F	df(error)
V	36.18*	2 (54)	2.34	4 (54)
A	12. 83*	2 (54)	1.77	4 (54)

*Note: * - p<0.01*

Table 5: The mean and standard deviation (SD) of peak amplitude (μV) across three stimulus conditions for three participant groups

Wave	Group	Condition		
		1	2	3
		Mean in μV (SD)	Mean in μV (SD)	Mean in μV (SD)
V	NML	0.17 (0.05)	0.10 (0.08)	0.13 (0.05)
	ML	0.15 (0.06)	0.13 (0.06)	0.13 (0.07)
	MC	0.15 (0.06)	0.14 (0.05)	0.12 (0.12)
A	NML	-0.13 (0.06)	-0.14 (0.07)	-0.12 (0.07)
	ML	-0.17 (0.06)	-0.14 (0.05)	-0.14 (0.07)
	MC	-0.19 (0.07)	-0.14 (0.08)	-0.21 (0.09)
C	NML	-0.23 (0.29)	-0.23 (0.29)	-0.16 (0.09)
	ML	-0.07 (0.02)	-0.09 (0.04)	-0.07 (0.04)
	MC	-0.25 (0.31)	-0.29 (0.50)	-0.20 (0.35)

Note:The waves A and C were recorded in the negative polarity and hence, the peak amplitude have a negative sign

Table 6: The results of repeated measure ANOVA for wave V and A amplitude

Wave	Effect of condition		Condition X Group	
	F	df(error)	F	df(error)
V	2.61	2 (54)	0.94	4 (54)
A	1.57	2 (54)	2.36	4 (54)

done using Bonferroni Post-hoc test. The results of the Post-hoc analysis demonstrated that the mean latencies were significantly prolonged in condition 2 and 3 compared to condition 1. There was no significant difference between condition 2 and 3, in their mean latencies. This was true for wave V as well as wave A. The representative waveform showing transient response comparison in three stimulus conditions is shown in Figure 3.

Descriptive statistics was done to obtain mean and standard deviation of peak amplitude in the three stimulus conditions (Table 5). Although mean amplitude differ across three stimulus conditions, there was no definable trend in the way mean amplitude of wave A varied among the three stimulus conditions. The peak amplitude was higher for stimulus condition 1 than that in condition 2 and 3 for wave V. The mean amplitude of wave V and A were compared across the three conditions using repeated measure ANOVA to verify the statistical significance of the observed differences. The results (Table 6) showed that there was no significant main effect of condition on amplitude of transient responses. Also, there was no significant interaction between group and condition.

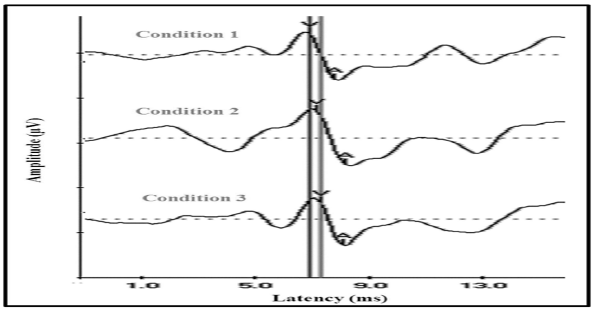


Figure 3: Transient portion of the waveforms recorded in the three stimulus conditions in an individual participant.

Effect of Condition on Sustained Responses

The peak amplitude at the frequencies corresponding to fundamental frequency (F0), first formant (F1) and, high frequency region (HF) of the stimuli was obtained from the FFT analysis (Figure 3). As apparent from Table 7, there is no general trend of the peak amplitude across the three stimulus conditions. The significance of difference in the mean amplitudes of F0, F1 and, HF was tested using repeated measure ANOVA taking group as a between-subject variable. The mean differences were however found to be statistically insignificant for F0 [F (2, 54) = 0.302, p > 0.05], F1 [F (2, 54) = 0.103, p > 0.05] and, HF [F (2, 54) = 1.069, p > 0.05].

Effect of Group on Speech Perception in Noise

The mean and standard deviation of the SNR-50 was obtained for the three groups. The results are shown in Figure 4. As seen in the figure, SNR-50 was lowest (better) for the musician group followed by music listeners and non music listeners. The mean differences among the three groups were compared on one-way ANOVA

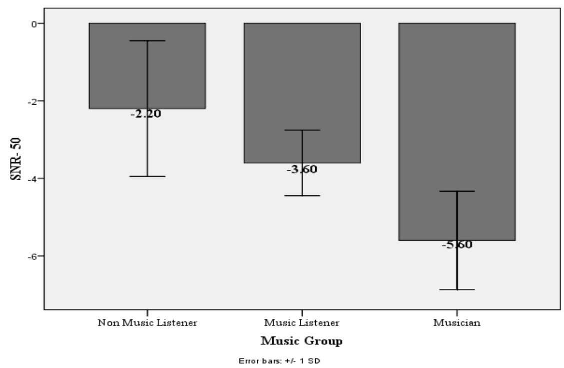


Figure 4: Graphical representation of the mean and standard deviation (SD) of SPIN in the three participant groups.

Table 7: The results of one-way ANOVA showing the effect of groups on latency of wave V and A

Wave	Condition	Effect of Group	
		F	df (error)
V	1	0.70	2 (27)
	2	1.94	2 (27)
	3	2.52	2 (27)
A	1	0.47	2 (27)
	2	0.26	2 (27)
	3	2.37	2 (27)

Table 8: The results of one-way ANOVA showing the effect of groups on amplitude of wave V and, A

Wave	Condition	Effect of Group	
		F	df (error)
V	1	0.51	2 (27)
	2	0.99	2 (27)
	3	0.08	2 (27)
A	1	2.34	2 (27)
	2	0.01	2 (27)
	3	3.48	2 (27)

taking group as an independent variable. The results revealed that there was significant main effect of group on SNR-50 [F (2, 27) = 16.289, p = 0.000]. Consequently, pair- wise comparison was done on Bonferroni post-hoc test. The results showed that the musician group had significantly better (lower) SNR-50 compared to the other two groups. There was no significant difference between mean SNR-50 of music listeners and non music listeners.

Effect of Group on Speech Evoked Brainstem Responses

When compared among the three participant groups, for wave V and A, musicians showed shorter latencies than music listeners, which in turn were shorter than the non music listener group. The mean amplitude of wave A was higher for the musician group compared to the other two groups. However, no such trend was seen in the mean amplitudes of wave V. To derive the group effect, the mean data were compared across the three groups on one-way ANOVA. This was done separately for each stimulus condition. The results of ANOVA (Table 7 & 8) showed that the group effect was absent on the latencies and amplitudes of transient response in all the three conditions.

The mean amplitudes of the F0, F1 and, HF (Table 9) were also compared across the three participant groups to study the effect of group on sustained responses. As evident from Table 9, mean was higher in musicians in contrast with music listeners and non music listeners for F0 and HF in condition 1, and F1 in all stimulus conditions. However, no such trend was seen for other responses. One-way ANOVA was done for the same and the results showed that the mean amplitude across the groups were not significantly different ($p > 0.05$). The F and degree of freedom (df) for each parameter in each condition are given in Table 10.

Effect of Group on Online Plasticity

The online plasticity was quantified by subtracting latencies and amplitude obtained in repetitive paradigm with that of latency and amplitude obtained in oddball paradigms using white noise. The resultant was termed as index of online plasticity. This was separately done for the data of each participant group. The mean and standard deviations of latency and amplitude index of online plasticity is given in Figure 5 and Figure 6, respectively. The mean results evidently show that these differences were smaller in musicians compared to non-musicians and, music listeners for amplitude index of online plasticity (except peak A). However, these differences were found to be statistically insignificant, on one-way ANOVA (Table 11).

The index of online plasticity was also computed by subtracting amplitude of sustained responses in repet-

Table 9: The mean and standard deviation (in parenthesis) of the amplitude (μV) of sustained responses across the three stimulus conditions for the three participant groups

Response	Group	Condition (μV)		
		1	2	3
		mean (SD)	mean (SD)	mean (SD)
F0	NML	5.61 (2.04)	5.42 (2.61)	6.73 (2.39)
	ML	5.49 (2.55)	6.47 (2.46)	5.83 (1.74)
	MC	6.98 (2.53)	5.48 (2.85)	6.11 (2.04)
F1	NML	0.66 (0.29)	0.65 (0.29)	0.70 (0.23)
	ML	0.65 (0.20)	0.64 (0.17)	0.67 (0.12)
	MC	0.88 (0.42)	0.87 (0.50)	0.79 (0.50)
HF	NML	0.30 (0.07)	0.32 (0.11)	0.34 (0.11)
	ML	0.31 (0.07)	0.29 (0.07)	0.31 (0.04)
	MC	0.34 (0.05)	0.31 (0.09)	0.33 (0.06)

Note: NML- Non Music Listener, ML- Music Listener, MC- Musician

Table 10: The results of one-way ANOVA showing the effect of group on FFR measures

Response	Condition	Effect of Group	
		F	df (error)
F0	1	1.20	2 (27)
	2	0.49	2 (27)
	3	0.49	2 (27)
F1	1	1.57	2 (27)
	2	1.37	2 (27)
	3	0.41	2 (27)
HF	1	1.17	2 (27)
	2	0.37	2 (27)
	3	0.39	2 (27)

Table 11: Results of one-way ANOVA showing the effect of group on online plasticity index

Measure	Wave	Effect of Group	
		F	df (error)
Latency	V	2.29	2 (27)
	A	0.64	2 (27)
Amplitude	V	1.85	2 (27)
	A	1.86	2 (27)

Table 12: The main effect of group for amplitude of FFR responses

FFR	Effect of Group	
	F	df (error)
F0	2.34	2 (27)
F1	0.01	2 (27)
HF	1.11	2 (27)

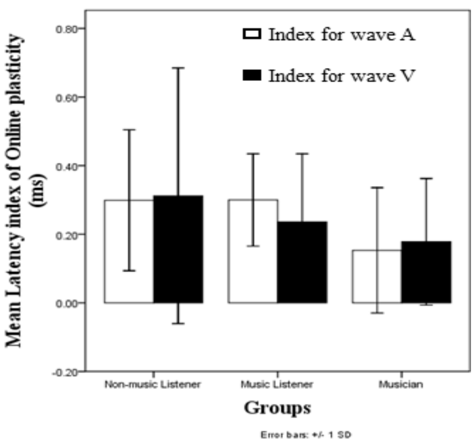


Figure 5: Graphical representation of the mean and standard deviation of online plasticity derived from latency of transient responses in the three participant groups.

itive paradigm from the odd-ball paradigm using white noise as context. The mean and standard deviations

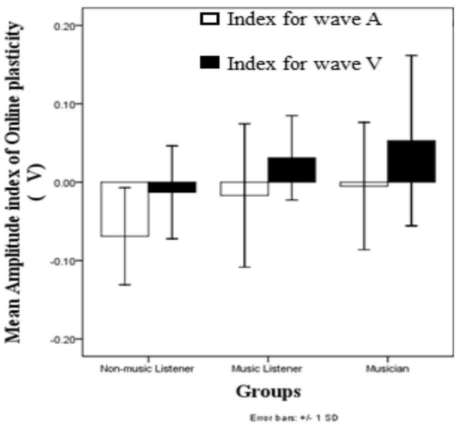


Figure 6: Graphical representation of the mean and standard deviation of online plasticity derived from amplitude of transient responses in the three participant groups.

are represented in Figure 7 (A to C). The mean amplitude differences of sustained responses were found to be consistently lower in musician group, when compared against non-music listener and, music listener group. One-way ANOVA however showed no significant variation in the amplitude values across groups ($p > 0.05$). The F and degree of freedom (df) for each parameter are given in Table 12.

Correlation of Music Training and SPIN

The analysis of group effect on SPIN showed that speech perception in noise in musicians was better compared to other two groups. Hence, it was of interest to study the relation between the years of training and SPIN scores. Figure 8 represent the scatter plot depicting the relation between SNR-50 and years of music training. The data of the two variables (SNR-50 &, years of training) in musician group were correlated using Pearson correlation. However, no correlation was found between the two variables ($r = -0.06$, $p > 0.05$).

Correlation between Online Plasticity and SPIN

The correlation between the SPIN performance and the index of online plasticity was established using Pearson correlation. The results showed that there was a positive moderate correlation between the SNR-50 and the wave V latency index of online plasticity ($r = 0.479$, $p < 0.01$). However, no correlation ($p > 0.05$) was found on the wave A latency index of online plasticity. This relation between online plasticity and SPIN is shown in Figure 9.

Discussion

Based on the knowledge about the mechanisms of training related neural plasticity and previous research reports, it is logical to assume that musicians have trained corticofugal pathway. The corticofugal pathway has been found to be moderating a newly proposed plas-

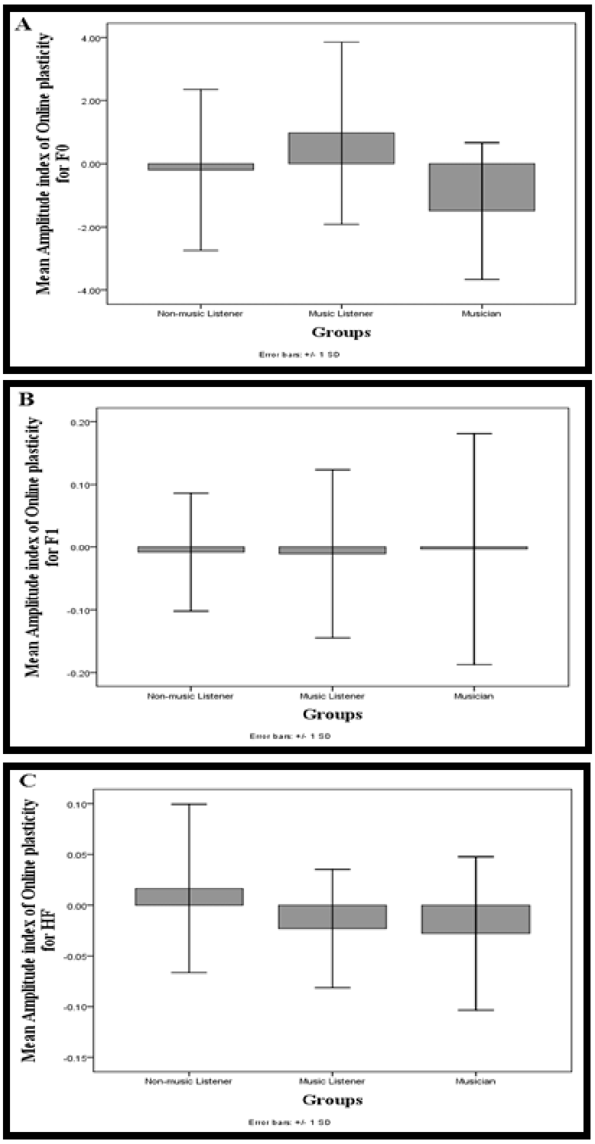


Figure 7: Graphical representation of mean and standard deviation (SD) of amplitude index of online plasticity for A) F0, B) F1, and C) HF across participant groups.

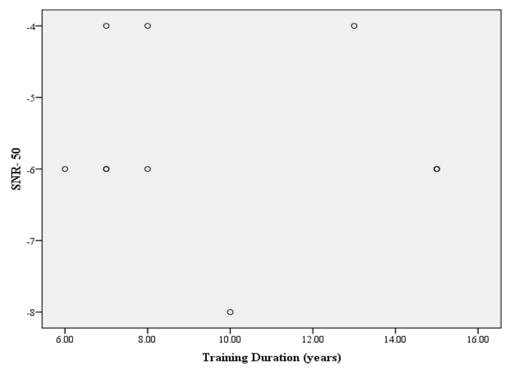


Figure 8: Relation between SNR-50 and years of music training.

ticity called online plasticity, which in turn is functional in enhancing speech perception in noise. In the present study, it was hypothesized that trained musicians have

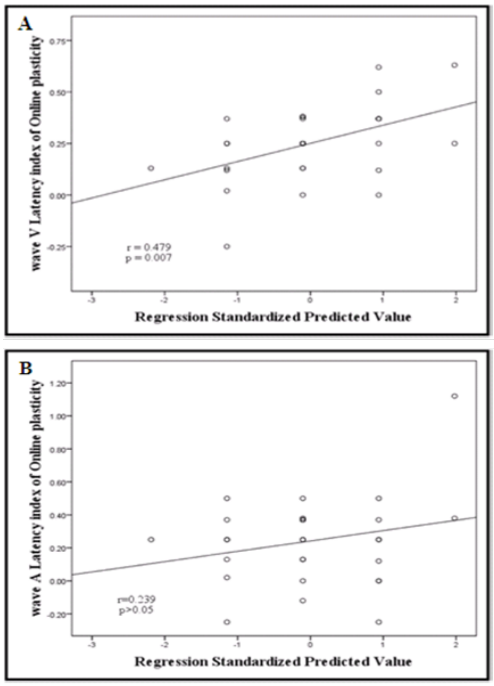


Figure 9: The correlation between Online plasticity and Speech perception in noise A) wave V latency and B) wave A latency.

better online plasticity and speech perception in noise compared to non-musicians. To further understand its mechanisms the online plasticity was compared among non-music listeners, music listeners and, musicians.

The wave C was identifiable in 100% of the musicians, which was not true for non-musicians (NMLs & MLs). Earlier studies have shown on electrophysiological studies that the encoding of speech is better among musicians when compared to age matched non-musician groups (Kraus, Skoe, Parbery-Clark & Ashley, 2010; Musacchia et al., 2007; Wong et al., 2007). These findings further strengthen the notion of neural synchrony being the determining factor for the occurrence of wave C. The lower occurrence of wave C in non-musician groups hence may be justified through reduced neural synchrony. The finding is also supported by the trend observed in the mean amplitude of wave C which was higher in musician group compared to non-musician groups.

The findings that the responses were better in the repetitive paradigm compared to that in the odd-ball paradigm is in consonance with the earlier findings (Chandrasekaran et al., 2009; Hanan & Maruthy, 2011). These are indicative of enhanced coding of the speech stimulus at the level of brainstem, when the stimulus is presented repetitively. This relative enhancement in the brainstem responses consequent to the repeating stimulus has been termed online plasticity (Chandrasekaran et al., 2009). As both the stimuli used as context (white noise & F2 filtered /da/) in the present study differed in the spectrum compared to the target /da/, the findings that both the contexts induced similar change further supports that it is the spectral difference that cues for context-dependent encoding and supports the inference of Hanan and Maruthy (2011).

The present finding of delayed transient responses in the contextual encoding is in contradiction to the earlier reports by Chandresekaran et al. (2009). Chandrasekaran et al. had seen a significant change only in the HF amplitude but not in the transient responses. Although the exact reason for the difference in the two studies is not clear, the present finding can be justified through the course of efferent pathway. The efferent pathway consists of multiple feedback loop system, which helps in the brainstem modulation. It is suggested that these feedback loop system selectively enhances relevant information in the signal, inhibiting the irrelevant information (Gao & Suga, 1998; Yan & Suga, 1998; Luo, Wang, Kashani & Yan, 2008). The result also duplicates the findings of Hanan and Maruthy (2011) who used the same paradigm. These findings are preliminary electrophysiological evidence for the corticofugal modulation of transient responses which may have implications for the perception of consonantal cues.

In Chandrasekaran et al. (2009), the context dependent effect on FFR was found at discreet intermediate frequencies (H2 & H4), while the effect was absent at F0, H3, H5 and H6. The absence of the context dependent effects in FFR in the present study may be because the analysis was over a wider range of frequencies (F0, F1, & HF), due to which the effect at some of the discreet frequencies might have got nullified.

The enhanced SPIN in musicians could be due to single or multiple underlying mechanisms (pertaining to afferent & efferent auditory pathway). The differences exist with respect to anatomical differences (Gaser & Schlaug, 2003; Ozturk, Tascioglu, Aktekin, Kurtoglu & Erden, 2002; Hyde et al, 2009; Bengtsson et al., 2005) and, enhanced encoding of spectral and temporal cues (Kraus & Chandrasekaran, 2010) in musicians when compared with non-musicians. The efferent pathway shows generation of the templates as a result of continuous representation of the ongoing stimulus (Haenschel, Vernon, Dwivedi, Gruzelier & Baldeweg, 2005; Strait, Kraus, Parbery-Clark & Ashley, 2010; Parbery-Clark, Strait & Kraus, 2011). These templates are especially essential for the exclusion of noise thus enhancing speech perception in noise (Chandrasekaran & Kraus, 2009). SPIN, in the past, also has been reported to be regulated by the OCB (Kumar & Vanaja, 2004). Deriving evidences from OCB studies (Micheyl, Khalfa, Perrot & Collet., 1997; Perrot, Micheyl, Khalfa & Collet, 1999; Ameen & Maruthy, 2011; Micheyl, Carbonnel & Collet, 2002) it could be concluded that the olivocochlear pathway is stronger in musician group compared to the non-musicians. Kumar, Hegde and Mayaleela (2010) provided evidence for changes in corticofugal modulation of olivocochlear bundle after short-term perceptual learning of non-native speech contrast. Probably, the enhanced speech in noise of musicians observed in the present study is a consequence of similar change in the olivocochlear bundle but, due to long-term formal musical training. Skoe, Banai and Kraus (2012) suggest that even the cognitive skills are essential for the improved perception of speech in noise. Thus, the present finding of better speech perception in noise in musicians can be justified through training related changes in the afferent auditory pathway, efferent auditory pathway or in cognitive domain.

The finding that there is no relationship between the number of years of music training taken and the speech perception in noise is in agreement with the findings of Parbery-Clark, Skoe, Lam and Kraus (2009). In the present study, the criterion to categorize participants into musician group was minimum of five years of formal music training. Based on these results, it could be inferred that the changes in the efferent system as a consequence of musical training would take place by five years of training.

The findings that there was no difference in the speech ABRs of the three groups in any of the conditions, means that music training as a variable does not influence brainstem encoding of speech. The results are in contrast with the earlier reports that the subcortical tuning is enhanced in the musicians compared to non-musician group, as evident in speech evoked ABR (Lee et al., 2009; Musacchia et al., 2007; Strait et al., 2009; Hyde et al., 2009; Hannon & Trainor, 2008). It has been reported that the anatomical and physiological changes are more evident in individuals with early-age (less than 7 years of age) of music training (Schlaug, Jancke, Huang, Staiger & Steinmetz, 1995; Pantev et al., 1998; Watanabe, Savion-Lemieux and Penhune, 2007). It was further concluded by Bailey and Penhune (2010) and, Penhune (2011) that there exist a sensitivity period, during which if musical training is given, would cause long-term improvement in the maturation of the pathway responsible for the sensorimotor integration. This could be the reason of the contrasting results obtained in the present study, as the mean onset age of musical training was 11;1 years as opposed to less than 5 years in other studies. The current results draw further support from the findings of Strait et al. (2009), where they reported that the musicians show distinct results from the non-musicians when compared with respect to the age of onset of music training or number of years of music training and, not when musicians with early-onset and late onset of training were grouped together.

The absence of group effect on FFT can also be attributed to the differences in methods of FFR analysis. In the previous studies the amplitude on FFT output (Lee et al., 2009; Musacchia et al., 2009) was measured at discreet frequencies. However, the analysis in the present study was over two ranges of frequencies. The method had been adopted from earlier publication (King et al., 2002; Wible et al., 2004; Werff & Burns, 2011).

The results of the experiment suggested that the amount of online plasticity is comparable among non-music listeners, music listeners and, musicians. This means that the music training or music listening did not influence the online plasticity as measured in the current electrophysiological paradigm. However, the conclusion is restricted to the group of musicians who started their training after about 11 years of age. Further, the finding also supports that the enhanced speech perception in noise observed in the musicians of this study is not related to the online plasticity of the brainstem. The present study showed a low positive correlation between online plasticity derived from wave V latencies and speech perception in noise. That means speech perception in noise improves with online plasticity. However, the relationship is not a strong one. This could be because; the speech perception in noise is determined by multiple factors like OCB functioning, binaural integration, working memory etc, and the influence of corticofugal pathway is only one of those factors. Hence, it can be inferred that to objectively study the correlates of behavioral speech perception in noise, one must study the online plasticity, OCB functioning, and binaural integration using physiological tests.

Conclusions

Overall, from the findings of the present study, it can be concluded that musicians who start their formal training after about 10 years of age do not have enhanced online plasticity. Online plasticity can be reliably documented using context-dependent encoding and is functional as it regulates speech perception in noise. The findings also demonstrated that only active tasks like singing and playing a musical instrument is advantageous for corticofugal regulation of speech perception in noise, not the relatively passive task like listening to music.

The study helps in understanding mechanisms of online plasticity and its role in speech perception in noise. It guides the audiologist in setting a protocol for evaluating context-dependent encoding of brainstem responses. It also guides clinical audiologists in the assessment of speech perception in noise and, understanding probable reasons for its deficits. Based on these findings audiologist can recommend music training to individuals with deficits in speech perception in noise.

References

American National Standards Institute. (1991). *American National Standard Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms.*ANSI S3.1- (1991). New York: American National Standards Institute.

Ameen, Md., & Maruthy, S. (2011). *Effect of music on neural plasticity of efferent auditory system.* Unpublished dissertation, University of Mysore, India.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural Timing Is Linked to Speech Perception in Noise. *The Journal of Neuroscience, 30*(14), 4922- 4926.

Anderson, S., & Kraus, N. (2011). Neural encoding of speech and music: Implications for hearing speech in noise. *Seminars in Hearing, 32,* 129-141.

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010). Brainstem correlates of speech-in-noise perception in children. *Hearing Research, 270* 151-157. doi:10.1016/j.heares.2010.08.001.

Bailey, J., & Penhune, V. (2010) Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. *Experimental Brain, Research, 204*, 91e101.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience, 8*(9), 1148-1150.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2010). Cross-domain Effects of Music and Language Experiencee on the Representation of Pitch in the Human Auditory Brainstem. *Journal of Cognitive Neuroscience, 23*(2), 425- 434.

Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology, 47,* 236-246.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009).Context dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron, 64,* 311-319.

Gao, E., & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency may in bat auditory system. *Proceedings of the National Academy of Sciences, USA, 95,* 12663-70.

Gaser, C., & Schlaug, G. (2003). Brain structure differ between musicians and nonmusicians. *The Journal of Neuroscience, 23*(27), 9240-9245.

Haenschel, C., Vernon, J. D., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event- Related Brain Potential Correlates of Human Auditory Sensory Memory-Trace Formation. *Journal of Neuroscience, 25,* 10494-10501.

Hanan, T., & Maruthy, S. (2011). *Effect of context on brainstem encoding of speech*. Unpublished dissertation. University of Mysore, Mysore.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical Training Shapes Structural Brain Development. *The Journal of Neuroscience, 29*(10), 3019 -3025.

Johnson, K., Nicol, T., Zecker, S., & Kraus, N. (2007) Auditory brainstem correlates of perceptual timing deficits. *Journal of Cognitive Neuroscience 19,* 376-385.

Johnson, K. L., Nicol, T. G., Zecker, S. G., Bradlow, A., Skoe, E., & Kraus, N.(2008). Brainstem encoding of voiced consonant-vowel stop syllables. *Clinical Neurophysiology, 119,* 2623-2635.

Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical "what" and "where" pathways in the auditory system. *Trends in Neurosciences, 28,* 176-181.

Kraus, N., McGee, T. J., & Carrell, T. D. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science, 273,* 971-973.

Kraus, N., Skoe,E., Parebery-Clark, A., & Ashley, R. (2009). Experience-induced Malleability in Neural Encoding of Pitch, Timbre, and Timing: Implications for Language and Music. *The Neuroscience and Music III- Disorders and Plasticity: Annals of the New York Academy of Sciences*, 1169, 543-557.

Krishnan, A. (2007). *Frequency-Following Response*. In R. F. Burkard, J. J. Eggermont, M. Don (Eds.). Auditory Evoked Potentials: Basic Principles and Clinical Application (pp. 313-335). Philadelphia, PA: Lippincott Williams & Wilkins.

Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain & Language, 110,* 135-148.

Krishnan, A., Swaminathan, J., & Gandour, J. T. (2008). Experience-dependent enhancement of Linguistic Pitch Representation in the Brainstem Is Not Specific to a Speech Context. *Journal of Cognitive Neuroscience, 21*(6), 1092-1105.

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research, 1313,* 124-133.

Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport, 20*(4), 408-413.

Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequencyfollowing response: Representation of pitch contours in Chinese tones. *Hearing Research, 189,* 1-12.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25,* 161-168.

Kumar, U. A., Hegde, M., & Mayaleela. (2010). Perceptual learning of non-native speech contrast and functioning of the olivocochlear bundle. *International Journal of Audiology, 49*(7), 488-496.

Kumar, U. A., & Vanaja, C. S. (2004). Functioning of Olivocochlear Bundle and Speech Perception in Noise. *Ear and Hearing, 25*(2), 142-146.

Luo, F.,Wang, Q., Kashani, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience, 28*, 11615-21.

Madhok, P., & Maruthy, S. (2010). *Neurophysiological consequence of auditory training: Subcortical and cortical structures.* Unpublished dissertation, University of Mysore, India.

Micheyl, C., Carbonnel, O. & Collet, L. (1995). Medial olivocochlear system and loudness adaptation: differences between musicians and nonmusicians, *Brain Cognition,* 29, 127-136.

Micheyl, C., Khalfa, S., Perrot, X., & Collet, L. (1997). Difference in cochlear efferent activity between musicians and nonmusicians. *NeuroReport, 8,* 1047-1050.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences, USA, 104*(40), 15894-15898.

Ozturk, A.H., Tascioglu, B., Aktekin, M., Kurtoglu, Z., & Erden, I. (2002). Morphometric comparison of the human corpus callosum in professional musicians and non-musicians by using in vivo magnetic resonance imaging. *Journal of Neuroradiology 29*, 29 -34.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E. & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392,* 811-13.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *The Journal of Neuroscience, 29*(45), 14100 -14107.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech in noise. *Ear and Hearing, 30*(6), 653-661.

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011).

Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia, 49,* 3338-3345.

Parbery-Clark, A., Strait D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical Experience and the Aging Auditory System: Implications for Cognitive Abilities and Hearing Speech in Noise. *PLoSONE, 6*(5), e18082. doi:10.1371/journal.pone.0018082.

Penhune, V. B. (2011). Sensitive periods in human development: Evidence from musical training. *Cortex, 47,* 1126-37.

Perrot, X., Micheyl, C., Khalfa, S., & Collet, L. (1999). Stronger bilateral efferent influences on cochlear biomechanical activity in musicians than in nonmusicians. *Neuroscience Letters, 262,* 167-170.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115*(9), 2021-2030.

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research, 156,* 95-103.

Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neurophysiologia, 33*(8), 1047-1055.

Shinn-Cunningham, B. (2001). Models of Plasticity in Spatial Auditory Processing. *Audiology and Neuro-Otology, 6*(4), 187-191.

Skoe, E., & Kraus, N.(2010b). Hearing it again and again: On-line subcortical plasticity in humans. *PlosONE.* 5(10): e13645, doi:10.1371/journal.pone.0013645.

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neuralefficiency- effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience*, 29, 661-668.

Strait, D.L., Kraus, N., Parbery-Clark, A.,, & Ashley R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing Research, 261,* 22-29.

Swaminathan, J., Krishnan, A., Gandour, J. T., & Xu (2008). Applications of Static and Dynamic Iterated Rippled Noise to Evaluate Pitch Encoding in the Human Auditory Brainstem. *IEEE Transactions on Biomedical Engineering, 55*(1), 281- 287.

Thakur, A. & Kumar, A. U. (2008). *Development of Quick speech in noise test for English speaking Indian children.* Unpublished dissertation, Mangalore University, Mangalore.

Tremblay, K., & Kraus, N. (2002). Beyond the ear: central auditory plasticity. *Otorinolaringology, 52,* 93-100.

Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America, 102*(6).

Watanabe, D., Savion-Lemieux, T., & Penhune, V. (2007). The effect of early musical training on adult motor performance: evidence for a sensitive period in motor learning. *Exp Brain Res, 176, 2,* 332-340.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*(4), 420-422.

Yan, W., & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience, 1,* 54-58.

Development of High Frequency Speech Identification Test in Manipuri Language

[1]Margaret Hmangte & [2]Geetha C.

Abstract

The present study aimed at developing and administering high frequency speech identification test in Manipuri language. The study consisted of two stages. In stage I, monosyllabic words, which majorly contained k/, /k^h/, /h/, /s/, /p/, /p^h/, /t/, /t^h/, /tʃ/ & /ʃ/ consonants and /i/, /e/ & /ei/ vowels, were selected from different sources. Long-term average speech spectrum was done on the selected words, to confirm that the selected words had high frequency spectral energy. Words with peak frequency of 2 kHz or above, and which had energy present even in the frequencies above the peak frequency were selected. After this, familiarity check was done. Only familiar and most familiar words were selected and two subtests of monosyllabic high frequency words were developed using those words. The developed materials were administered at three presentation levels [20, 40 and 60 dB SL (Ref: SRT)] to 20 normal hearing individuals at each level. SIS was obtained using both the subtests. The results showed a significant difference between the three levels, i.e., with increase in the levels, the SIS increased, with almost 100% at 40 dB SL. Comparison across the two subtests showed that any of the two subtests can be used to obtain high frequency speech identification scores. Thus, from the above findings, it may be concluded that this test has got normal performance-intensity function and has list equality. However, the clinical utility of the test has to be assessed by administering it on clinical population with different degrees of hearing impairment.

Keywords:*Speech Identification Scores, monosyllables, Long-Term-Average Speech Spectrum.*

Introduction

A person with a hearing loss is bound to have difficulty in the perception of speech sounds. Depending on the pattern of audiogram, the speech perception ability of an individual varies. A most common audiogram configuration, which results in a poor speech perception, is high frequency sloping hearing loss.

Individuals with sloping high frequency hearing loss would have difficulty mainly in speech sounds having energy concentration in the high frequency region (Dean & Mc Dermott, 2000; Risberg & Martony, 1972). Experiments have repeatedly shown that speech understanding cannot be predicted from pure tone thresholds, especially in these cases. Young and Gibbons (1962) noted that although there are some degrees of association between scores obtained from test of speech understanding and pure tone thresholds, the relationship was not strong enough to allow accurate prediction of speech understanding from pure tone audiogram. This is especially true for individuals whose hearing loss is not of flat pattern. Hence, carrying out speech audiometry for the assessment of auditory recognition or identification of words, nonsense syllables or phonemes, is a necessary part of clinical evaluations for individuals with hearing impairment.

Speech Identification Score (SIS), which is a part of speech audiometry, gives information on the speech identification ability of the person. For obtaining SIS, normally, a test which contains all consonants and vowels of the language is used. However, such a test would not give a correct picture about the identification ability of a person with a high frequency sloping hearing loss. This is because the person would perceive the low and mid frequency sounds relatively better and thus, would get clues from these frequencies. This will result in a better SIS than the person's actual speech identification in real life situations.

Hence, it is important to have the test material which contains only high frequency consonants and vowels present in that particular language, when assessing individuals with high frequency hearing loss.

The first high frequency word list was developed by Gardner (1971) in English language. Gardner developed a word list that contained consonants of high frequency spectral energy and used it for testing speech discrimination in cases of high frequency hearing loss. It has also been found to be helpful in the application of hearing aid selection and auditory training.

There are also other high frequency word lists available in English language. They are; Pascoe High Frequency Test (Pascoe, 1975); and California Consonant Test (Owens & Schubert, 1977).

It is a well-known fact that an individual's perception of speech is reported to be influenced by his/her mother tongue (Singh & Black, 1966). Most people consistently had better and optimum discrimination scores in their mother tongue as compared to other languages (De, 1973). Hence, speech materials have been developed in Indian languages for evaluating Indian population.

[1]Email: hmangtemargaret@gmail.com,
[2]Lecturer in Audiology, Email: geethamysore.cs@gmail.com

The Indian subcontinent consists of a number of separate linguistic communities each of which share a common language and culture. The people of India speak many languages and dialects. It is important to have speech materials developed in each of these languages. There are high frequency word lists developed and standardized in some of the Indian languages such as High Frequency-Kannada Speech Identification Tests (Kavitha, 2002), High Frequency Speech Identification Test in Tamil (Sinthiya, 2009), and High Frequency Speech Identification Test in Telugu (Ratnakar, 2010).

Manipuri is one of the official languages of India. It is spoken by over 56% of population in Manipur. Manipuri has several high frequency consonants (/t̪/, /t̪^{h}/, /s/, /dZ/ /tʃ/, /tʃ^h/ and /ʃ/) and high frequency vowels (/i/, /e/ and /ei/). Further, the occurrence of these high frequency sounds is not uncommon in Manipuri language.

Speech materials to assess Speech Recognition Scores and Speech Identification Test developed by Devi (1985) are in Manipuri language. However, the Speech Identification Test contains all the sounds present in the language.

As mentioned earlier, Speech Identification Test containing all the sounds in the language have low, mid and high frequencies, which make the list redundant for a person with high frequency sloping hearing loss. Hence, it does not give true communication difficulties of persons with high frequency hearing loss.

Further, for the selection of appropriate hearing aids for individuals with sloping hearing loss, it is essential that a test that is sensitive to their problems be utilized. Therefore, speech identification scores obtained for a high frequency word list is a better means to assess the individual with high frequency hearing loss.

In addition, with increase in number of geriatrics, who most often exhibit sloping type of loss, there is an urgent need for developing High Frequency Speech materials in Manipuri language for assessing speech perception of the individuals who speak Manipuri language. Therefore, the present study attempts to develop a High Frequency Speech Identification Test in Manipuri Language.

Hence, the aim of the present study was to develop High Frequency Speech Identification Test in Manipuri language and to administer the developed test on participants with normal hearing sensitivity who are native speakers of Manipuri, at three different input levels, i.e., 20 dB SL, 40 dB SL and 60 dB SL (Ref. SRT).

Method

The objectives of the study were to develop High Frequency Speech Identification Test in Manipuri language, and to administer this on participants with normal hearing sensitivity at different input levels. The study was conducted in two stages. Stage 1 involved development of high frequency word subtests in Manipuri language. Stage 2 involved administering the developed test on participants with normal hearing sensitivity.

Stage 1: Development of the High Frequency Word Subtest in Manipuri Language

The following steps were involved, for preparing the high frequency test material in Manipuri language: Selection of words with high frequency sounds in Manipuri language; Recording of the selected words; Analyzing Long-term average speech spectrum (LTASS) of the selected words; Familiarity assessment; and Construction of subtests.

Selection of words with high frequency sounds in Manipuri language: Monosyllabic words, 230 in number, were selected from different sources like newspapers, dictionary, text books etc. The monosyllabic words, which majorly contained k/, /k^h/, /h/, /s/, /p/, /p^h/, /t/, /t^h/, /tʃ/ and /ʃ/ consonants and /i/, /e/ and /ei/ vowels were chosen. In the literature, these phonemes have been reported to have energy at high frequencies and thus, results in confusion for individuals with high frequency hearing loss (Cooper, Liebermann, Delattre, Brost & Gerstmann, 1952; Gardner, 1971; Hughes & Halle, 1956).

Recording of the Test Material: The recording of these 230 words were done in a sound treated double room. The monosyllabic words were spoken by an adult female who was native speaker of Manipuri language. This was recorded into a computer using 16 kHz sampling rate and 16 bit quantization using Computerized Speech Lab (CSL) 4500 software.

The speaker was instructed to say the words with flat tone and to keep the loudness constant across the words. The VU meter was monitored within optimum levels during the recording. The signal was digitized at a sampling rate of 16 kHz using the 12 bit analog to digital and digital to analog converter housed within a computer. Noise and hiss reduction was carried out on the recorded materials and amplitude normalization of the signals was done using the Adobe Audition (version 3.0) software to maintain constant amplitude across the words. The recorded materials were played to two adults who were native speakers of Manipuri language to ensure that the articulation and the clarity of the recorded material were good.

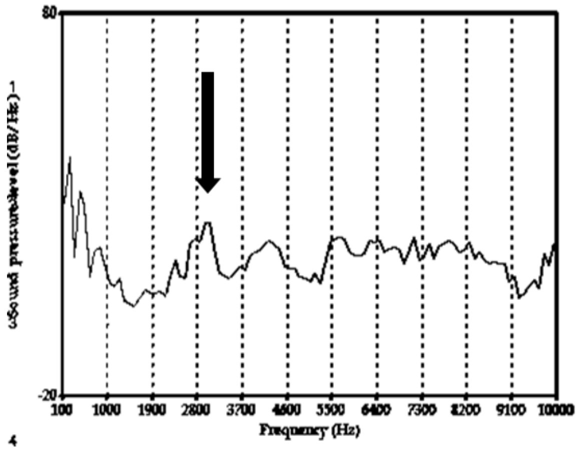


Figure 1: Spectrum of the word /siŋ/. Long Term Average Speech Spectrum showing peak frequency at 2800 Hz (indicated with an arrow mark) and the presence of energy even above 2800 Hz.

Long-Term Average Speech Spectrum (LTASS) of the selected words: LTASS was done on this 230 words, to confirm that the selected words have high frequency spectral energy. LTASS was derived using PRAAT (version 4.1) software and the spectral information was determined manually. The peak frequency of the spectrum was taken as the target parameter. Words with peak frequency of 2 kHz or above, and which had energy present even in the frequencies above the peak frequency were selected.

Familiarity Assessment: Familiarity assessment of the words selected based on the LTASS were assessed by five adults, two males and three females, who were native speakers of Manipuri language. They were instructed to rate for familiarity of each word on a three point scale; 1) most familiar (the words which were commonly used by the individual), 2) familiar (the words that are used occasionally) and 3) unfamiliar (the words which are not used). The words that were rated as most familiar and familiar were selected for the construction of the two subtests with one list each.

Construction of Word Subtest: Two subtests with 25 words each were made after familiarization. The frequencies of occurrence of sounds were balanced as far as possible between the two subtests. These are given in Appendix.

The audio recorded files of these words were then copied to a compact disk. The inter stimulus interval between the words was set to 3 seconds. A calibration tone of 1 kHz was inserted before beginning of the high frequency word lists to adjust the VU meter at zero.

Stage 2: Administration of the Test Material on Participants with Normal Hearing Sensitivity

Selection of Participants: In this stage, the developed high frequency test materials were administered on 60 native speakers of Manipuri. For the selection of the participants for the study, routine audiological evaluation was carried out. The individual was considered for the study if the person was a native Manipuri speaker within the age range of 18 to 30 years. The person had to have normal hearing sensitivity (PTA less than or equal to 15 dB HL) from 250 Hz to 8000 Hz for air conduction and 250 Hz to 4000 Hz for bone conduction. They had normal middle ear condition in both the ears, with 'A' or 'As' type of tympanogram with ipsilateral and contralateral reflexes present at 500, 1000 and 2000 Hz. The person no history of otological (ear disease, trauma, ototoxic drug intake or ear operation) or neurological dysfunction and had normal speech and language skills.

Testing Environment: All the evaluations were carried out in an acoustically treated two-room situation. This set up had minimum noise levels.

Instrumentation: A calibrated dual channel GSI 61 audiometer coupled with acoustically matched TDH 39 headphones housed in MX-41AR ear cushions and B71 bone vibrator was utilized to estimate the pure tone threshold, speech recognition thresholds (SRT) and speech identification scores (SIS). Calibrated GSI Tympstar middle ear analyzer was used for obtaining tympanogram and acoustic reflex thresholds.

A computer was used to present the recorded speech test material. The output was routed through a computer connected to the auxiliary input of the calibrated GSI 61audiometer through TDH 39 headphones.

Test Procedure: The pure tone thresholds were tracked for frequencies between 250 Hz to 8 kHz for air conduction and 250 Hz to 4 kHz for bone conduction using the modified Hughson and Westlake procedure (Carhart & Jerger, 1959). The speech recognition threshold and SIS were obtained using the existing SRT and SIS lists

Table 1:List of 78 words with their peak frequency

Sl. No.	Words	LTASS	Sl. No.	Words	LTASS	SL. No.	Words	LTASS	SL. No	Words	LTASS
1.	/ki/	2000	21.	/cæŋ/	2800	41.	/t̪ʰɛt/	2500	61.	/sɛp/	2800
2.	/ke/	2800	22.	/cɛt̪/	2900	42.	/t̪ʰem/	2800	62.	/sɛm/	2800
3.	/ken/	2800	23.	/cen/	2900	43.	/t̪ʰəi/	2800	63.	/sku:l/	3700
4.	/kəi/	2800	24.	/cep/	2900	44.	/pi/	2000	64.	/sɤit/	2500
5.	/kʰi/	3500	25.	/cəi/	2900	45.	/pik/	2800	65.	/slet̪/	2000
6.	/kʰik/	2500	26.	/ṭik/	2000	46.	/pin/	2800	66.	/svəi/	2000
7.	/kʰiŋ/	2500	27.	/ṭiŋ/	2800	47.	/pɛk/	2000	67.	/svi/	2200
8.	kʰin/	2500	28.	/t̪in/	2000	48.	/pəi/	2000	68.	/hi/	2000
9.	/kʰɛɤ/	2200	29.	/t̪e/	2000	49.	/pʰi/	2000	69.	/hik/	2000
10.	/kʰe/	2500	30.	/t̪ɛk/	2000	50.	/pʰəi/	2000	70.	/hiŋ/	2500
11.	/kʰɛk/	2500	31.	/t̪ɛŋ/	2000	51.	/si/	2800	71.	/hip/	2500
12.	/kʰɛt/	2500	32.	/t̪en/	2000	52.	/sik/	2800	72.	/hui/	2000
13.	/kʰəi/	2000	33.	/t̪ɛm/	2000	53.	/sit/	2800	73.	/hɛn/	2000
14.	/ci/	2500	34.	/t̪əi/	2000	54.	/siŋ/	2800	74.	/hɛk/	2000
15.	/cik/	2500	35.	/tʰi/	2800	55.	/sin/	2800	75.	/həi/	2000
16.	/ciŋ/	2000	36.	/t̪ʰiŋ/	2800	56.	/sɛ/	2300	76.	/dzo:i/	2800
17.	/cit/	3000	37.	/t̪ʰit/	2800	57.	/sɛk/	2000	77.	/Zip/	2000
18.	/cin/	2000	38.	/t̪ʰin/	2800	58.	/sɛŋ/	2800	78.	/dzəil/	2800
19.	/cɛ/	2200	39.	/t̪ʰɛk/	2200	59.	/sɛt/	2800			
20.	/cɛk/	2900	40.	/t̪ʰɛŋ/	2000	60.	/sɛn/	2800			

in Manipuri language developed by Tanuja (1985). A GSI Tymstar middle ear analyzer was used to find out the type of tympanogram and acoustic reflexes at 500 Hz, 1 kHz and 2 kHz.

Administration of the Test Material: Prior to testing, external input to the audiometer was calibrated to 0 VU, using a 1000 Hz calibration tone, for each participant. The high frequency speech identification lists developed were played through CD player at 20 dB SL, 40 dB SL and 60 dB SL (Ref. SRT).

The participants were asked to listen to the instructions first and to follow the instructions. Stimuli were presented through headphones and an open set response in the form of an oral response was obtained. All participants were tested monaurally using the developed lists. The tester recorded the responses in a scoring sheet.

Scoring: Word scoring was done for both the lists. Scoring was done by giving a score of '1' for a correct repetition and '0' for a wrong repetition or missed words.

Statistical Analysis: The Statistical Package for the Social Sciences (version 17.0) software was used to carry out the statistical analysis. Repeated measures ANOVA, and Bonferroni pairwise comparison test were carried out for the analysis of the data.

Results and Discussion

The aim of the present study was to develop a High Frequency Speech Identification Test in Manipuri language and to administer this on participants with normal hearing sensitivity at three levels of presentation.

Development of the high frequency word lists

As mentioned in the chapter 3, monosyllabic words with the phonemes (/h/, /s/, /p/, /p^h/, /t/, /t^h/, /tʃ/, /ʃ/, /k/ & /k ^{h}/), 230 in number, were selected from different sources like newspapers, dictionary, text books etc. These 230 words were recorded by two female adult native speakers of Manipuri. Recording was done in a sound treated room by using 16 kHz sampling rate and 16 bit quantization using Computerized Speech Lab (CSL) 4500 software. After recording, LTASS was done for these words to assess the spectrum.

Results of LTASS

In order to confirm that these words have energy at high frequencies, LTASS was done using PRAAT (version 4.1) software. From the spectrum, the peak frequency and the pattern with which the energy spreads above the peak frequency were analyzed. The words which were found to have peak energy at 2 kHz or above, and with energy present even above the peak frequency were selected.

Out of 230 words, 78 words met the above mentioned criteria. The list of these words and the frequency at which peak energy was present is given in the Table 1. The selected words were then assessed for familiarity by five adult native speakers of Manipuri language.

According to Gardner (1971), consonants /p/, /t/, /k/, /s/, /f/, / θ/, /h/) result in confusion for individuals with high frequency hearing loss. In addition, even the fricative /ʃ/ had energy above 2000 to 3000 Hz frequency (Hughes & Halle, 1956) and the affricate /tʃ/ is also included in the high frequency consonants. Hence, it is important to include all these consonants along with vowels (Cooper, Liebermann, Delattre, Brost & Gertmann, 1952).

Further, LTASS has been successfully used in the present study to ensure that the energy in low frequencies is not dominating and confirm the high frequency energy spread of the selected words. This has also been used in several other studies for the development of speech materials (Kavitha, 2002; Sudipta, 2006; Sinthyia, 2009; Ratnakar, 2010). Out of these 78 words, 60 words were rated as most familiar and familiar, and thus, were selected for preparing the lists. These words were not sufficient to make three lists of 25 words each, and hence, 50 words with highest peak frequency were selected for preparation of two subtests with 25 words each. These 50 words are highlighted in the Table 1 and the two subtests are given in appendix.

Comparison of SIS across levels within and between the lists

Two subtest with 25 words each were administered on 60 normal individuals at three levels i.e. 20 dB SL, 40 dB SL and 60 dB SL on 20 subjects at each levels. Hence, there were three groups of participants. Since the number of sample was lesser in each group, to ensure normal distribution of the sample, One- sample Kolmogorov - Smirnov test was carried out. The result of this test is given in the Table 2. The test results showed that the standard level is minimal for all the three groups. Hence, the sample tested is normally distributed.

Table 2: Results of one sample test (maximum score=25)

	Mean(Number of correctly identified words)	SD	Z score	p- value
20 dB SL	21.6	1.79	1.19	0.11
40 dB SL	24.6	0.77	2.91	0.00
60 dB SL	25.0	0.00	-	-

Table 3: Mean and Standard Deviation for SIS for two subtests across three presentation levels

Subtests	20 dBSL Mean(SD)	40 dBSL Mean(SD)	60 dBSL Mean(SD)
1	21.95 (1.93)	24.70 (0.57)	25.00 (0.00)
2	21.25 (1.61)	24.55 (0.94)	25.00 (0.00)

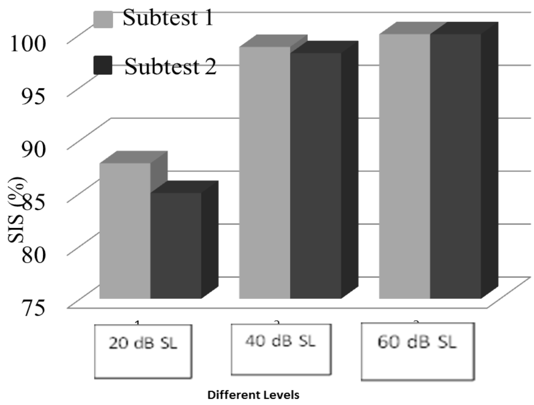


Figure 2: Figure 2: Comparison of SIS across levels for both the subtests.

Comparison between the SIS obtained at three different levels for both the subtests were made. The mean and SD of SIS at different levels for the two subtests are given in the Table 3. It can be observed that the SIS at 20 dB SL is at around 92%, at 40 dB SL, it is around 99%, and 100% at the highest level of presentation, i.e., 60 dB SL. In order to see if there is a statistically significant difference across two subtests and three different levels repeated measures ANOVA was carried out. Results of repeated measures ANOVA revealed that there was a significant difference between the SIS at three different levels [F (2, 76) = 127.901, $p < 0.05$].

As there was a significant difference between the SIS obtained at three levels, Bonferoni pair-wise comparison of SIS was done across different presentation levels.

The results showed a significant difference across three levels ($p < 0.01$). That is, the scores at 20 dB SL were poorer and were significantly different from the SIS obtained at 40 dB SL and 60 dB SL. The scores at 40 dB SL were a little lesser and were significantly different from the SIS obtained at 60 dB SL, i.e., as the level of presentation increased the scores also increased.

The result of the present study is consistent with the results of Turner and Cummings (1999). They studied the SIS across different input levels, i.e., performance-intensity function, in the normal participants of 20 - 27 years of age. They found that with the increase in intensity the scores improved, and also reported that the scores were near 100% at the level of 50 dB SPL in normal hearing participants. In the present study, also, there was near 100% performance (as can be seen in Figure 4.3) for both the subtests, at 40 dB SL (which is around 70 dB SPL), which is higher than that was reported in the above study. There could be two explanations for this. One could be that the testing was not done at 30 dB SL which would be almost 60 dB SPL. The second reason for this could be that, in the above study, the task was closed set easier task. In the present study, it was an open set task, which is relatively difficult. Further, Sinthiya (2009) also found near 100% responses at 40 dB SL (ref. SRT) in normal hearing subjects.

Comparison of SIS between the two subtests was also done (given in the table 3). It can be observed that there is not much of difference in the scores between the two subtests. Results of repeated measures ANOVA showed no significant difference between the SIS obtained for the two subtests [$F(1, 38) = 1.494$, $p > 0.05$], and there was no significant interaction effect found between the subtests and the different presentation levels [$F(2, 76) = 1.250$, $p > 0.05$]. This indicates that performance obtained from the two subtests will yield similar results. Hence, it can be said that the two subtests developed in the present study have good equality.

Conclusions

It can be concluded that this test has got normal performance-intensity function and has list equality. However, the clinical utility of the test has to be assessed by administering it on clinical population with different degrees of hearing impairment.

References

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

Cooper, F. S., Delattre, P. C., Libermann, A. M., Borst, J. M., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America, 24,* 597-606.

Dean, M. R., & McDermott, H. J. (2000). Speech perception with steeply sloping hearing loss, effects of frequency transposition. *British Journal of Audiology, 34,* 353-361.

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

Gardner, H. J. (1971). Application of high frequency consonant discrimination word list in hearing aid evaluation. *Journal of Speech and Hearing Disorder, 36*(3),354-355.

Hughes, G. W., & Halle, M. (1956).Spectral properties of fricatives consonants. *Journal of the Acoustical Society of America, 28,* 303-310.

Jerger, J., & Jerger, S. (1971). Diagnostic significance of PB word functions. *Archives of Otolaryngology, 93,* 573-580.

Kavitha, E. M. (2002). *High Frequency-Kannada Speech Identification Test.* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore.

Levin, R. A. (1952). *The intelligibility of different kinds of test material used in speech audiometry.* Master's Dissertation, University of Washington.

Maroonroge, S., & Diefendorf. (1984).Comparing normal hearing and hearing impaired individuals subjects performance on North-Western Auditory test No.6, California Consonant Test & Pascoe's High frequency word list. *Ear and Hearing, 5*(6), 356-360.

Mendel, L. L., & Danhauer, J.L. (1997). *Audiological evaluation and management and speech perception assessment.* San Diego. London: Singular Publishing Group Inc.

Owens, E., & Schubert, E. D. (1977). Development of California Consonant Test. *Journal of Speech and Hearing Research, 20,* 463-474.

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,* 1-40.

Ratnakar, Y. V. (2010). *High Frequency Speech Identification in Telugu.* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore.

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.

Sinthiya, K. (2009). *High Frequency Speech Identification in Tamil.* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore.

Singh, S., & Black, J. W. (1966). Study on 26 intervocalic consonants as spoken and recognised by four language groups. *Journal of the Acoustical Society of America, 39,* 372-387.

Sudipta, K. B. (2006). *High Frequency-English Speech Identification Test (HF- ESIT).* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore.

Devi, T. E. (1985). *Development and standardization of speech test materials in Manipuri language.* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore.

Turner, C. W., & Cummings, K. J. (1999).Speech Audibility for Listeners with High-Frequency Hearing Loss. *American Journal of Audiology, 8,* 47-56.

Young, M. A., & Gibbons, E. W. (1962).Speech discrimination scores and threshold measurement in a non-normal hearing population. *Journal of Audiological Research, 2,* 21-33.

Effect of Spectrally and Temporally Modulated Maskers on Speech Perception in Listeners with Auditory Dys-Synchrony, Cochlear Hearing Loss and Normal Hearing

[1]Merry Elizabeth Roy & [2]Animesh Barman

Abstract

The present study aimed at assessing the speech recognition performance in individuals with auditory dys-synchrony (AD), cochlear hearing loss and normal hearing in the presence of spectrally and temporally modulated maskers at 0 dB and 10 dB SNR and to observe which clinical group would take greater advantage of spectral and/or temporal dips to understand speech. Number of words correctly identified within each sentence was calculated in the presence of each of the modulated (spectrally modulated noise with 4 ERB gaps, spectrally modulated noise with 2 ERB gaps and temporally modulated noise) and unmodulated (speech shaped steady state noise) maskers at both SNRs in 10 individuals with AD, 13 individuals with cochlear hearing loss and 20 individuals with normal hearing sensitivity. All three groups performed poorer at 0 dB SNR than at 10 dB SNR and in the presence of unmodulated than modulated maskers. The AD group performed significantly poorer under temporally modulated noise at both the SNRs, while they showed better performance when masker was spectrally modulated. This could be attributed to the excessive masking in them due to smearing of the temporal waveform. The Cochlear hearing loss group did not benefit from 2 ERB gap spectral modulation and temporal modulation of noise. The AD group performed significantly poorer compared to the other two groups on all conditions due to their affected temporal resolution, while this group also showed significant release from masking for spectrally modulated maskers compared to the other groups.

Keywords: *Unmodulated masker, spectrally modulated masker, temporally modulated masker*

Introduction

Speech is considered to be a complex dynamic signal which fluctuates both in amplitude and frequency over time. To perceive these inherent fluctuations in the signal, the auditory system does a detailed spectral and temporal analysis of the signal. Normal perception is hence directly dependent on an intact peripheral and central auditory processing. But the perception of speech is intricate when distorted or attenuated in the presence of noise. This difficulty in perception seen even in normal hearing individuals is yet more unfavorable in those individuals with hearing impairment.

Studies (Festen & Plomp, 1990; Moore, 1996) have reported that listeners with normal hearing and those with hearing impairment have difficulty in the perception of speech in noisy and reverberant conditions. This is because noise reduces the redundancy that is available inherently within the signal. As the noise dominates, i.e., the speech to noise ratio (SNR) reduces, it becomes more difficult to understand speech. But if the noise or the background sound also fluctuates in time, there are moments or dips created where the speech is distinctive of noise. Individuals with normal hearing have the ability to recognize speech with much accuracy in such fluctuating backgrounds than in steady state or continuous noise (Festen & Plomp, 1990) unlike those with hearing impairment (Peters, Moore & Baer, 1998).

The modulated or fluctuating maskers are characterized by spectral and temporal dips. The temporal dips are instants when the overall level of the background noise is low during which the signal-to-noise ratio is high, which allows brief 'glimpses' to be obtained of the target speech. The spectral dips arise when the spectrum of the target speech signal over any short interval is different from that of the background noise. Although some parts of the target speech spectrum may be completely masked by the background, other portions of the signal during periods in which the masker reaches a dip is utilized to infer the complete target speech. This benefit received when listening to speech in the presence of fluctuating maskers than in the presence of steady state maskers is referred to as 'release of masking'.

However, studies have reported that individuals with cochlear hearing loss do not show this benefit, i.e., they perform almost similarly in presence of modulated and steady-state maskers (Middelweerd, Festen, & Plomp, 1990; Festen & Plomp, 1990).

Duquesnoy (1983) measured the speech recognition threshold (SRT) required to correctly identify 50% of the stimuli in presence of amplitude modulated noise and showed that a difference in SRT ranging from about 7dB to 15 dB exists between individuals with normal hearing and those with cochlear hearing loss. Peters et al. (1998) reported that, SRTs decreased by only 1 to 2 dB when the bandwidth of spectral dips of the masker was increased from two to four ERBNs in hearing impaired listeners, whereas SRTs decreased by 6 dB

[1]Email: merryroy88@yahoo.com,
[2]Reader in Audiology, Email: nishiprerna@yahoo.com

for normal hearing listeners in comparison to a steady state masker. This reduced ability to take the benefit of spectral and temporal dips seen in these individuals with cochlear hearing loss could be attributed to the reduced temporal and spectral resolution (Wagener, Brand & Kollmeier, 2006; Peters, et al., 1998).

The potential to hear low-level speech segments and to resolve spectral dips is largely determined by the active mechanism in the cochlea, which depends on the functioning of outer hair cells (Moore, 2003). But in case of cochlear hearing loss, all the three factors important for release from masking: audibility, spectral resolution, and temporal resolution may be adversely affected (Moore, 2007).

Bernstein and Grant (2009) proposed that the magnitude of masking release also depends on the signal-to-noise ratio (SNR) at which performance is measured i.e., release from masking tends to be large when the SNR is low, and small or absent when the SNR is high. This means that it is important to compare the performances of hearing impaired and normal hearing listeners at different SNRs.

Analogous to those having cochlear hearing loss, individuals with Auditory Dys-synchrony (AD) have also shown to be having reduced spectral (Kraus et al., 2000) and temporal processing (Zeng, Kong, Michalewski & Starr, 2005). Rance, McKay and Grayden (2004) found significant correlation between reduced speech perception abilities and extremely poor temporal processing and frequency discrimination ability.

These deficits could be attributed to the reduced synchrony in neural firing which disrupts the timing cues and affects the listener's ability to cope with the dynamic nature of speech signals. It could impair not only the ability to use amplitude envelope cues in speech, but also to perceive rapidly changing spectra in the speech stimuli (Rance et al., 2004). Individuals with AD, are known to exhibit even greater difficulty for perceiving speech in the presence of noise. Kraus et al. (2000) have reported that individuals with AD, obtain significantly depressed scores in the presence of a multi talker speech babble, in spite of performing remarkably well in quiet.

Zeng and Liu (2006) reported that even at SNRs that show little or no effect on individuals with normal hearing (10 to 15 dB), these individuals show detrimental scores which is supported by psychophysical studies showing excessive masking effects in them (Zeng et al., 2005; Zeng, Oba & Starr, 2001; Kraus et al., 2000). The mechanisms underlying excessive noise effects in AD type hearing loss are unclear, although there is psychophysical evidence that auditory signals are more affected by simultaneous and non-simultaneous masking than normal listeners in these individuals with AD (Vinay & Moore, 2007; Zeng et al., 2005; Kraus et al., 2000).

Recent studies suggested that neural phase locking to the temporal fine structure of the target signal may be critical for listening in the background temporal dips (Moore, Glasberg & Hopkins, 2006; Leger, Moore & Lorenzi, 2012). It may thus be presumed that the reduced phase locking ability in these individuals with AD may hinder release from masking.

Considering the natural conditions of speech perception in background noises that are temporally and spectrally varying, such as clattering dishes or background conversations, the investigation of speech perception in the presence of fluctuating or modulated backgrounds is important. Intact spectral and temporal resolution in individuals with normal hearing sensitivity allows them to utilize the spectral and temporal dips in noise. Psychophysical studies have pointed out that individuals with cochlear hearing loss and those with AD exhibit spectral and temporal resolution problems. Thus a comprehensive knowledge about psychophysical findings reported in literature could be better corresponded with the speech perception difficulties. The present study was hence undertaken to examine the effects of maskers which are modulated either spectrally or temporally on speech perception in individuals with normal hearing, cochlear hearing loss and auditory dys synchrony.

The study aimed to assess speech recognition performance in groups of individuals AD, cochlear hearing loss and normal hearing in presence of spectrally and temporally modulated noise at 0 dB SNR and 10 dB SNR. It also aimed to observe which clinical group would take greater advantage of spectral and/or temporal dips to understand speech.

Method

Participants

To accomplish the goal, a total of 43 participants participated in the study. They were categorized into 3 groups. Group I consisted of 10 individuals with AD in the age range 18 to 55 years having pure tone thresholds within 55 dB HL with either flat or gradually rising audiogram. Group II included 13 Individuals with cochlear hearing loss of age ranging from 18 to 55 years with pure-tone thresholds between 25 to 55 dB HL having a flat audiometric configuration. Group III consisted of 20 normal hearing listeners age matched with that of individuals in group I and II of age ranging from 18 to 55 years

The actual experiment was carried out in two phases.

Phase 1: Preparation of the Stimulus

Target speech stimuli: Seven lists of sentences were taken from standardized quick SIN test in Kannada developed by Methi, Avinash and Kumar, (2009) to assess the speech recognition ability in the participants from all the three groups. Each list contains 7 sentences and each sentence has 5 key words, making a total of 35 keywords in each list.

Maskers: Following are the ipsilateral maskers and the procedure to generate, used to determine the SIS:

Speech shaped steady state noise: A Speech shaped noise or SSN was generated from the whole set of sentences at a sampling frequency of 44.1-kHz by estimating the long-term power spectrum of recorded test sentences. This was done by randomizing the phase of the Fourier spectrum of concatenated words of original signals using MATLAB (version 2009). It had a spectrum which approximates the average long term spectrum of the target sentences spoken by an adult male with a secondary peak present around 100 Hz.

Speech shaped noise with spectral modulations: The speech shaped steady state noise was filtered so as to have spectral dips in several frequency regions. The filtering was done based on the equivalent-rectangular-bandwidth (ERB) scale derived from the auditory filter bandwidths for normally hearing participants (Glasberg & Moore, 1990). The relationship between the number of ERBs and frequency is,

ERB number = 21.4 $\log_{10}$ (4.37*F*+1).

Each ERB represents one auditory filter bandwidth.

The noise was filtered in 2 ways as shown in the Figure 1: first, with an alternating pattern of two ERBs present and two ERBs removed (spectrally modulated noise with 2 ERB gaps) and the second with an alternating pattern of four ERBs present and four ERBs removed (spectrally modulated noise with 4 ERB gaps)

Speech shaped noise with temporal modulations: Speech spectrum-shaped wide-band noise was modified to have envelope modulations or temporal fluctuations imposed on it. This was achieved by modulating the amplitude of speech shaped noise at the rate of 10 Hz using MATLAB software (version 2009). This noise was referred to as 'temporally modulated noise'.

The rms level of all these noises were adjusted according to the level of the target speech stimuli to achieve the desired SNR using MATLAB software (version 2009). The noises were mixed with the passages using MATLAB software at 2 different SNRs. A total of 7 conditions were prepared using 7 sentence lists to assess sentence perception at two SNRs. The conditions at both 0dB SNR and 10 dB SNR included the following noise types: speech shaped noise, spectrally modulated noise with 4 ERB gaps, spectrally modulated noise with 2 ERB gaps and temporally modulated noise. An

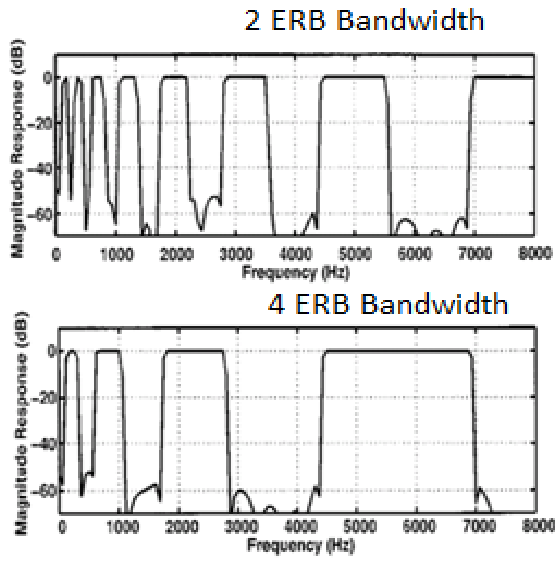


Figure 1: Characteristics of the digital filters used to produce the noises with multiple spectral notches

additional testing condition in the presence of Speech shaped noise at 10 dB SNR was prepared for group I: individuals with AD. This was done based on the results of a pilot study revealing very poor scores at 0dB SNR for all noise conditions. Hence to make a better comparison of modulated and unmodulated masker conditions, this additional condition was prepared. Randomly selected sentences from list 1 and list 2 were mixed with speech shaped steady state noise at 10dB SNR which served as an additional testing condition for individuals with AD.

All the 7 lists of sentences were used for each of the 7 conditions mentioned above. Thus a total of 49 lists were made. These 49 lists were randomly grouped into 7 sets of sentence lists, such that each set had all the 7 test conditions. Hence, each participant was tested with all seven lists having 7 different conditions, so as to avoid any effect of a particular list on the performance. These 7 testing conditions were administered in a randomized order across participants and also within each list, sentences were presented in random.

The Adobe audition software (Version 3) was used to normalize the stimuli to a level of -15dB. The order of presentation followed the manner such that always lists with sentences at 0 dB SNR was presented before the sentences presented at 10 dB SNR. These prepared stimuli were transferred digitally to a recordable compact disc for use in the experiment. The CD had a total of 9 tracks. Track 1 had a calibration tone of 1 kHz with a level identical to the normalized level of the stimuli. Using the 1-kHz calibration tone, VU meter on the audiometer was adjusted to read '0'. Tracks 2-8 had 7 sets of lists with different stimulus conditions as mentioned in the order earlier. Track 9 had the additional lists prepared to administer on individuals with AD at 10 dB SNR with speech shaped steady state noise.

Phase 2: Determination of Speech Identification scores in the presence of ipsilateral maskers

The target sentences mixed with noises were presented at 40 dB SL. The speech recognition scores were determined in 7 different conditions mentioned earlier. An additional testing condition at +10 dB SNR in presence of speech shaped steady state noise was administered on individuals with AD.

The testing was done in a two room testing condition. The stimuli was replayed manually by a PC and was routed through a calibrated (ANSI, 1996) diagnostic audiometer (GSI-61). It was presented monaurally to the participants through TDH 50P headphones. Participants were told that they would hear sentences in quiet and in noisy background and they were instructed to repeat verbally or write down what they heard. Only one ear was considered for all the participants to avoid the practice effect. Preferably right ear was chosen, otherwise ear with better speech recognition scores was selected.

Scoring

Each testing condition had a list with 7 sentences having 5 keywords in each sentence. The speech identification for each condition was calculated by counting the number of words the participant correctly identified. Each of the correctly identified key word was awarded one point for a total possible score of 35 points per list for each condition. The number of correctly identified words obtained using speech shaped steady state noise at 0 dB SNR and 10 dB SNR provided a reference condition against which speech identification obtained in other types of noises with spectral and temporal dips were compared. And as a measure of release from masking, number of correctly identified words under unmodulated speech shaped noise was subtracted from the scores obtained for each of the modulated noise condition separately. This was done so as to compare the release obtained with each of the modulated noise condition at a specific SNR.

Results

The results obtained are presented under within and across group comparisons.

Within Group Comparisons

Individuals with AD

Mean and standard deviation of number of correctly identified words (WRS) obtained for various noise conditions in 10 individuals with AD was calculated and tabulated in the Table 1.

Table 1: Mean and SD of number of correctly identified words (WRS) obtained for various noise conditions in individuals with AD

Conditions		Mean (no. of words)	SD
10 dB SNR	SSN	25.40	6.68
	ERB2	27.90	4.81
	ERB4	30.60	5.46
	AM10	21.40	7.77
0 dB SNR	SSN	04.60	5.18
	ERB2	11.10	5.87
	ERB4	19.00	8.53
	AM0	11.60	8.47

The various noise conditions are expanded as follows.

10 dB SNR SSN: Speech shaped steady state noise (SSN) at 10 dB SNR

10 dB SNR ERB2: Spectrally modulated noise with 2 ERB gaps (ERB2) at 10 dB SNR

10 dB SNR ERB4: Spectrally modulated noise with 4 ERB gaps (ERB4) at 10 dB SNR

10 dB SNR AM10: Temporally modulated noise (AM10) at 10 dB SNR

0 dB SNR SSN: Speech shaped steady state noise (SSN) at 0 dB SNR

0 dB SNR ERB2: Spectrally modulated noise with 2 ERB gaps (ERB2) at 0 dB SNR

0 dB SNR ERB4: Spectrally modulated noise with 4 ERB gaps (ERB4) at 0 dB SNR

0 dB SNR AM0: Temporally modulated noise (AM0) at 0 dB SNR

Note: abbreviations are the same for the consecutive tables also.

From the table it can be noted that mean of number of correctly identified words (WRS) obtained at 10 dB SNR is higher than that obtained at 0 dB SNR. At both the SNRs, WRS obtained for modulated maskers are greater than that obtained for the un modulated masker.

Effect of various maskers on number of correctly identified words (WRS) at different SNRs: Repeated measure ANOVA was done to see the effect of various maskers at 2 different SNRs on number of correctly identified words (WRS) in individuals with AD. The results indicated a significant difference in the number of correctly identified words (WRS) across noise conditions [$F (3, 27) = 15.021$ $p< 0.001$] at 0 dB SNR. At 10 dB SNR also, there was a significant difference in number of correctly identified words obtained across noise conditions [$F (3, 27) = 6.360$, $p <0.01$]. Bonferroni's pairwise comparison was done to see in which two condi-

Table 2: Results of Bonferroni's pairwise comparison of scores obtained between noises at 0 dB SNR in group with AD

0 dB SNR	ERB2	ERB4	AM0
SSN	Significant p<0.05	Significant, p<0.01	Not significant, p>0.05
ERB2		Significant, p<0.01	Not significant, p>0.05
ERB4			Significant, p<0.05

Table 3: Results of Bonferroni's pairwise comparison of scores obtained between noises at 10 dB SNR in group with AD

10 dB SNR	ERB2	ERB4	AM10
SSN	Not Significant, p>0.05	Not Significant, p>0.05	Not significant, p>0.05
ERB2		Significant, p<0.01	Not significant p>0.05
ERB4			Significant, p<0.05

Table 4: Mean and standard deviation of difference in WRS due to release from masking obtained in individuals with AD

Conditions	Modulated- Unmodulated	Mean (WRS difference)	SD
10 dB SNR	ERB2 - SSN	2.50	7.05
	ERB4 - SSN	5.20	7.89
	AM10 - SSN	-4.00	7.43
0 dB SNR	ERB2 - SSN	6.50	5.33
	ERB4 - SSN	14.40	8.35
	AM0 - SSN	7.00	9.38

tions, the number of correctly identified words (WRS) obtained differed significantly, both at 0 dB SNR and +10 dB SNR. Details of Bonferroni's test results are shown in Table 2 for 0 dB SNR and Table 3 for +10 dB SNR respectively.

Amount of release from masking obtained (improvement in number of correctly identified words) under various modulated maskers at 0 dB and 10 dB SNR: Release from masking was calculated by subtracting the number of correctly identified words (WRS) obtained in presence of unmodulated noises from modulated noises at 0 dB SNR and 10 dB SNR separately. Release from masking was measured by subtracting the number of words correctly identified in the fluctuating masker condition (3 conditions) by the number of words correctly identified in the steady speech shaped noise masker condition.

The mean and standard deviation for amount of release from masking in terms of improvement or reduction in number of correctly identified words (WRS) were calculated. The details are shown in Table 4.

The mean value shows a greater release from masking when the noise is spectrally modulated with 4 ERB gaps than with 2 ERB gaps at both the SNRs. It can also be noted that, at 10 dB SNR, temporally modulated noise did not show any benefit, compared to a steady state noise. To see whether these effects are significant or not, one way repeated measure ANOVA was done. The results revealed that the amount of release obtained with all 3 modulated noise conditions were different and was statistically significant at both 0 dB SNR [$F(2, 18) = 12.954$, $p < 0.001$] and 10 dB SNR [$F(2, 18) = 11.097$, $p < 0.001$]. On Bonferrroni's pairwise comparison, the pattern of results obtained was same at both SNRs and details are as shown in the Table 5.

The Table 5 shows that there is a significant release from masking in terms of number of correctly identified words (WRS) in the presence of spectrally modulated noise with 4 ERB gaps, when compared to other modulated maskers at both the SNRs.

0.0.1 Individuals with cochlear hearing loss

The mean and standard deviation of number of correctly identified words (WRS) under the various types of noise was obtained for all 13 individuals with cochlear hearing loss and tabulated in Table 6.

The mean of number of correctly identified words (WRS) obtained for 10 dB SNR is higher than that obtained at 0 dB SNR. It was also noted that the number of correctly identified words (WRS) obtained in the presence of spectrally modulated noise having 4 ERB gaps, were almost equal at both the SNRs.

Effect of various maskers on number of correctly identified words (WRS) at different SNRs: One way repeated measure ANOVA was done to see the effect of various maskers at different SNRs on number of correctly identified words (WRS) in individuals with cochlear hearing loss. The results showed a significant difference across noise conditions [$F(3, 36) = 5.879$ $p < 0.01$] at 0 dB SNR. Bonferrroni's pairwise analysis revealed a significant difference in 3 comparisons as shown in Table 7.

Table 5: Results of Bonferroni's pairwise comparison of differences in WRS obtained for modulated and unmodulated noises at both 0 dB SNR and 10 dB SNR in individuals with AD

Modulated - Unmodulated	ERB4-SSN	AM-SSN/AM10-SSN
ERB2-SSN	Significant, p<0.01	Not significant, p>0.05
ERB4-SSN		Significant, p<0.01

Table 6: Mean and SD of WRS obtained for various noise conditions in individuals with cochlear hearing loss

Conditions		Mean (WRS)	SD
10 dB SNR	ERB2	34.38	0.96
	ERB4	34.69	0.85
	AM10	34.00	1.91
0 dB SNR	SSN	27.76	4.53
	ERB2	29.15	5.45
	ERB4	32.53	3.43
	AM0	28.00	6.39

Table 7: Results of Bonferroni's pairwise comparison of differences in WRS obtained for modulated and unmodulated noises at both 0 dB SNR and 10 dB SNR in individuals with AD

0 dB SNR	ERB2	ERB4	AM0
SSN	Not Significant p>0.05	Significant, p<0.001	Not significant p>0.05
ERB2		Significant p<0.05	Not significant p>0.05
ERB4			Significant p<0.05

Table 8: Mean and standard deviation for amount of release obtained (improvement in number of correctly identified words) with modulated noises in comparison to unmodulated noise in individuals with cochlear hearing loss

Conditions	Modulated minus Unmodulated	Mean (WRS difference)	SD
0 dB SNR	ERB2 - SSN	1.38	5.73
	ERB4 - SSN	4.76	3.13
	AM0 - SSN	.23	5.01

Table 9: Results of Bonferroni's pairwise comparison of amount of release from masking obtained (improvement in number of correctly identified words) at 0 dB SNR in individuals with cochlear hearing loss

Modulated minus Unmodulated	ERB4-SSN	AM0-SSN
ERB2-SSN	Significant, p<0.05	Not significant, p>0.05
ERB4-SSN		Significant, p<0.01

It is evident from the Table 7 that number of correctly identified words (WRS) in the presence of spectrally modulated masker with 4 ERB gaps was significantly more than any other conditions. However word identification did not differ significantly between un modulated masker and other types of modulated maskers. However, at 10 dB SNR, there was no significant difference across the noise conditions [F (2, 24) = 1.16, p >0.05].

Amount of release from masking obtained (improvement in number of correctly identified words) under various modulated maskers at 0 dB SNR: Amount of release from masking in terms of improvement in number of correctly identified words (WRS) was calculated by subtracting the WRS obtained in the presence of unmodulated noises from modulated noises at 0 dB SNR as done for the previous group. The amount of release was not obtained at 10 dB SNR, because in all conditions all the individuals obtained almost maximum WRS possible and a test condition of unmodulated masker at 10 dB SNR was not included in the experiment in this group for comparisons. Thus improvement in terms of number of correctly identified words (WRS) due to release from masking at 10 dB SNR could not be observed. Mean and standard deviation of improvement in correctly identified words (WRS) at 0 dB SNR are tabulated in Table 8.

The mean value shows a greater release from masking when the noise is spectrally modulated with 4 ERB gaps than with 2 ERB gaps. It was also noted that temporally modulated noise showed almost no release from masking. To see if these effects are statistically significant or not, one way repeated measure ANOVA was done to compare the release from masking with different noise conditions at 0 dB SNR. The results showed that all 3 modulated noise conditions are significantly different [F (2, 24) = 7.174, p< 0.01]. Bonferrroni's pairwise comparison revealed significant differences between 2 comparisons as shown in the Table 9.

The Table 9 shows that there is a significant release from masking in terms of improvement in number of correctly identified words (WRS) in presence of spectrally modulated noise with 4 ERB gaps over the other two modulated masker conditions at 0 dB SNR.

Individuals with normal hearing sensitivity

The mean and standard deviation of number of correctly identified words (WRS) obtained in 7 different conditions in 20 individuals with normal hearing sensitivity are given in the Table 10.

The mean of number of correctly identified words (WRS) obtained at 10 dB SNR showed a ceiling effect across all noise conditions, which restricted any further comparison across the conditions at 10 dB SNR.

Effect of various maskers on number of correctly identified words (WRS) at different SNRs: One way repeated

Table 10: Mean and SD of number of correctly identified words (WRS) obtained for various noise conditions in individuals with normal hearing sensitivity

Conditions		Mean (WRS)	SD
10 dB SNR	ERB2	35.00	0
	ERB4	35.00	0
	AM10	35.00	0
0 dB SNR	SSN	32.25	2.57
	ERB2	33.55	2.01
	ERB4	34.30	1.30
	AM0	33.75	2.07

measure ANOVA was done to see the effect of various maskers on number of correctly identified words at 0 dB SNR. The results revealed a significant difference across the noise conditions [F (3, 72) = 13.313 p< 0.001] at 0 dB SNR. Bonferrroni's pairwise analysis showed significant differences between 2 comparisons as seen in the Table 11.

Table 11: Results of Bonferroni's pairwise comparison of WRS between noises in individuals with normal hearing sensitivity

0 dB SNR	ERB2	ERB4	AM0
SSN	Significant, p<0.01	Significant, p<0.001	Significant, p<0.001
ERB2		Not Significant, p> 0.05	Not significant p>0.05
ERB4			Not significant p>0.05

The Table 11 revealed that individuals with normal hearing obtained significantly better WRS in presence of all types of modulated maskers when compared to the unmodulated masker.

Amount of release from masking obtained (improvement in number of correctly identified words) under various modulated maskers at 0 dB SNR: Amount of release from masking was calculated by subtracting the scores obtained in presence of unmodulated noises from modulated noises at 0 dB SNR. The amount of release was not obtained at 10 dB SNR, because in all conditions all the individuals obtained maximum WRS possible. Mean and standard deviation of improvement in number of correctly identified words (WRS) at 0 dB SNR are tabulated in Table 12.

Table 12: Mean and standard deviation of amount of release obtained (improvement in number of correctly identified words) with modulated noises at 0dBSNR in individuals with normal hearing sensitivity

Modulated- Unmodulated	Mean(WRS difference)	SD
ERB2 - SSN	1.30	1.94
ERB4 - SSN	2.05	1.98
AM0 - SSN	1.50	1.98

The mean value shows almost similar amount of release across all types of noises. One way repeated measure ANOVA was done to compare the release from masking with different noise conditions, at 0 dB SNR. The release obtained with all 3 modulated noise conditions were not different significantly [F (2, 38) = 2.048, p>0.05].

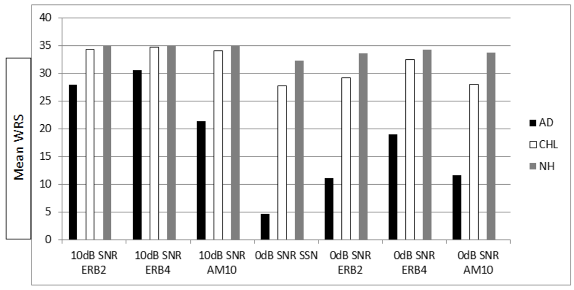


Figure 2: Mean of number of correctly identified words (WRS) obtained by three groups of participants across the various masking conditions.

Between Group Comparisons

Effects of different types of noise, group and SNR on number of correctly identified words (WRS)

Mean and standard deviation of number of correctly identified words (WRS) obtained for all the noise conditions at both SNRs in all three groups of participants are shown in the Figure 2.

From the figure, we can observe that all the three groups perform comparatively poorer at 0 dB SNR than at 10 dB SNR. Group with AD scored the least scores across all conditions compared to the other two groups. Individuals with normal hearing as well as those with cochlear hearing loss perform almost similarly at 10 dB SNR. All the groups scored poorer in unmodulated noise than compared to modulated noises. The amount of improvement in WRS for the modulated noise differed across the groups. Maximum scores were obtained in the condition where noise is spectrally modulated with 4 ERB gaps across all the groups at both SNRs.

Mixed ANOVA was done to see the main effects of groups, SNR and noises (3x2x3) (excluding the speech shaped steady state noise). The speech shaped steady state noise was excluded from overall comparison, because a masking condition with this noise at 10 dB SNR was not performed in groups with normal hearing sensitivity and cochlear hearing loss. The main effect of groups was highly significant [F (2, 40) = 69.061, p< 0.001]. The main effect of types of noises [F (2, 80) = 62.950, p< 0.001] and SNRs [F (1, 40) = 178.744, p < 0.001] were also highly significant. It was also found that there was a significant interaction between all the 3 variables: SNRs and groups [F (2, 40) = 54.317, p < 0.001]; noise and groups [F (4, 80) = 24.019, p< 0.001]; SNRs and noises [F (2, 80) = 6.341, p<0.01] and SNRs, noises and groups [F (4, 80) = 4.050, p<

0.01]. This indicates that performance in presence of modulated noises varied across groups.

Effect of types of noise on number of correctly identified words (WRS) irrespective of groups

Bonferroni's pairwise comparison was done to see if there are any significant differences in WRS between the noises, irrespective of the groups at each SNR, as Mixed ANOVA showed significant effect of different types of noise on word identification. At both SNRs, results followed a similar pattern which is shown in Table 13.

Table 13: Results of Bonferroni's pairwise comparison of WRS between types of noise at 0 dB and 10 dB SNR

0 dB SNR/10 dB SNR	ERB4	AM
ERB2	Significant $p<0.001$	Significant $p<0.01$
ERB4		Significant $p<0.001$

It was found that the 3 noises differed significantly from each other at 0 dB SNR and 10 dB SNR.

Effect of groups on WRS across noises at 0 dB SNR and 10 dB SNR:

To compare the scores obtained for four different noise conditions across the 3 groups at 0 dB SNR and 10 dB SNR, MANOVA was carried out. It was found that there was a highly significant ($p < 0.001$) difference between the groups across all the four noise conditions. Details are given in Table 14.

Table 14: F- values obtained across three groups at 0 dB SNR and 10 dB SNR for each of the noise conditions

Conditions		F values at $p < 0.001$
0 dB SNR	SSN	$F(2, 40) = 172.518$
	ERB2	$F(2, 40) = 92.455$
	ERB4	$F(2, 40) = 40.068$
	AM0	$F(2, 40) = 54.099$
10 dB SNR	ERB2	$F(2, 40) = 33.084$
	ERB4	$F(2, 40) = 10.192$
	AM10	$F(2, 40) = 45.761$

Duncan's post hoc test was done to see if the groups differed from each other for every noise condition at 0 dB SNR and 10 dB SNR. Duncan's post-hoc test ranked this difference in three homogeneous subsets for SSN, ERB2 and AM0 at 0 dB SNR. The results showed that at all conditions, group with AD differed significantly from the other two groups.

Amount of release from masking obtained (improvement in number of correctly identified words) across the groups at 0 dB SNR

Improvement in word identification due to release from masking in different groups was considered only at 0 dB SNR. It was not considered at 10 dB SNR, as groups having normal hearing group and cochlear hearing loss obtained maximum possible scores for all the conditions. The mean and SD values obtained at 0 dB SNR are shown in Table 15.

Table 15: Mean and standard deviation of amount of release obtained (improvement in terms of number of correctly identified words) in 3 groups of participants at 0 dB SNR

0 dB SNR	ERB2 - SSN	ERB4 - SSN	AM10 - SSN
AD	6.50 (5.33)	14.40 (8.35)	7.00 (9.38)
Cochlear HL	1.38 (5.73)	4.76 (3.13)	0.23 (5.01)
Normal Hearing	1.30 (1.94)	2.05 (1.98)	1.50 (1.98)

It can be observed that all groups showed a greater amount of release for spectrally modulated noise with 4 ERB gaps compared to other modulations in the noise. Groups with cochlear hearing loss and normal hearing sensitivity do not show much difference between them.

Mixed ANOVA was also done to see the overall effects of release from masking obtained with the three modulated noise conditions and to see the interaction between the release from masking and groups at 0 dB SNR. It was found that there was a significant main effect of amount of release from masking, across the modulated noises [$F(2, 80) = 31.033$, $p <0.001$]; across the groups [$F(2, 40) = 12.075$, $p <0.001$] and also a significant interaction between the amount of release and the groups was found [$F(4, 80) = 8.193$, $p <0.01$]. These results imply that the release may be different across different groups. Bonferroni's pairwise comparison was done to see if any significant difference exists between the amount of release obtained for each modulated noise, irrespective of the groups. The results are shown in Table 16.

Table 16: Bonferroni's pair wise comparisons for release obtained with modulated noises in comparison to unmodulated noise irrespective of groups

Noises at 0 dB SNR	ERB4-SSN	AM0-SSN
ERB2-SSN	Significant $p<0.001$	Not Significant $p>0.05$
ERB4-SSN		Significant $p<0.001$

The results revealed significant difference in amount of release obtained under spectrally modulated noise with 4 ERB gaps compared to other types of modulated noises irrespective of the groups.

To compare the amount of release obtained for 3 modulated noise conditions across the 3 groups at 0 dB SNR, MANOVA was carried out. It was found that there was a highly significant difference between the groups across all the comparisons as seen in Table 17.

Duncans post hoc test was done to see if the groups differed from each other in terms for amount of release from masking obtained with modulated maskers at 0 dB

Table 17: F values obtained across three groups at 0 dB SNR for 3 modulated noises

Conditions		F value
0 dB SNR	ERB2- SSN	F (2, 40) = 5.663, p<0.01
	ERB4 - SSN	F (2, 40) = 25.181, p<0.001
AM0 - SSN	F (2, 40) = 4.943, p<0.01	

SNR. It was found that group with auditory dys synchrony was significantly different from the other two groups in terms of release of masking obtained with modulated noises.

Discussion

The results obtained from various groups across various maskers at different SNRs on speech identification ability have been discussed below.

Individuals with AD

The results revealed that individuals with Auditory Dys-synchrony have significantly poorer speech identification scores in presence of speech shaped steady state noise (un modulated noise) and at 0dB SNR. This could be due to excessive masking effects in individuals with AD as reported by many authors (Zeng et al., 2005; Rance et al., 2004; Starr et al., 2003; Zeng et al., 1999). Zeng and Liu (2006) reported significantly poorer speech identification in presence of speech spectrum shaped noise at 0 dB SNR in 13 individuals with AD when compared to individuals having normal hearing and cochlear hearing impairment. Zeng et al. (1999) reported that impaired ability to follow temporal fluctuations in the signal is likely the underlying cause for the poor speech recognition in individuals with AD. A demyelinating neuropathy would lead to less faithful temporal representation of the signal due to loss of neural synchrony because; dys synchronous firing of neural impulses would reduce the number of neural spikes within each bin. Buss, Hall and Grose (2004) stated that individuals with AD are impaired in extracting both envelope and fine structure cues from speech signal and hence adding noise to the signal would exaggerate this difficulty. Physiologically, these excessive masking effects could be due to inner hair cell loss or loss of synchronous firing due to damaged nerve fibers (Harrison, 1998; Starr et al., 1996).

Demyelinated fibers may also display emphatic transmission (cross-talk) between fibers, with one active fiber cutting off discharges in adjacent fibers (Starr, Picton & Kim, 2001). This cross talk of fibers may lead to broader than normal neural tuning curves and this might lead to severe distortion in the coding of complex sounds like speech.

The current results pointed out that no significant differences in the speech identification was obtained in presence of speech shaped steady state noise and temporally modulated noise at 0 dB SNR as well as at 10 dB SNR. This indicates that these individuals with AD do not have the ability to take the advantage of temporal dips or modulations in noise, which could also be attributed to the poor temporal processing in these individuals (Zeng et al. 2005; Zeng et al. 1999). Eggermont (1997) stressed on the importance of neural synchrony across populations of neurons in the signaling of differences between a dynamic and a steady state signal. Hence the dys-synchronous neural discharge would have prevented these individuals with AD from detecting the temporal modulations in the signal. Figure 3 depicts the phenomenological model of auditory dys synchrony given by Zeng et al. (1999) to explain the temporal processing deficits in individuals with AD which could have led to poorer gap detection ability.

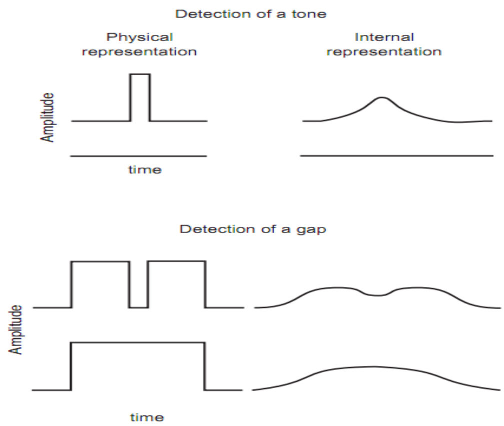


Figure 3: A phenomenological model of AD (Zeng et al., 1999). Desynchronous neural activity results in a smeared internal representation of a physical stimulus. Smearing of the temporal envelope does not affect the detection of a tone (top panel) because this task requires all or none decision. However, smearing causes greater problem in gap detection (bottom panel) as the task requires finer discrimination of two wave forms.

This abnormal smearing of the temporal waveform due to the dys-synchronous neural firing would fill in the temporal gaps in noise, thereby making the gaps unavailable for them to access glimpses of target speech. The persistence of effects of noise in the gaps could also be validated with the findings reporting excessive forward and backward masking in these individuals (Zeng et al., 2005), preventing them to separate out successive signals. The cross talk of nerve fibers which leads to broader neural tuning could also result in temporal smearing and hence poor detection of gaps in noise to perceive speech. Thus the average neural response to speech in presence of a temporally modulated background would be similar to the one in presence of un modulated signal.

In spite of these excessive masking effects, the current study found that individuals with AD are able to take the benefit of spectral modulations imposed on to the steady state speech spectrum shaped noise at 0 dB SNR which was statistically significant. This benefit was observed for both spectral modulations at 0 dB SNR, i.e. with 2 ERB gaps as well as with 4 ERB gaps. However, temporal modulation in noise also showed improvement in WRS, but was not significantly different from that obtained in presence of steady state noise. This implicated relatively intact spectral processing in individuals with AD enabling them to detect spectral gaps in noise. Psychophysical tuning curves in individuals with AD have shown sharper tips indicating normal OHC function (Vinay & Moore, 2007).

Near normal frequency discrimination ability at higher frequencies (Zeng et al., 2005) and normal auditory filter bandwidth (Rance et al., 2004) have also been reported in these individuals. Therefore it could be assumed that the intact detection of spectral gaps in noise allowed the high frequency information in the target speech to be perceived relatively better when compared to a steady state noise. However due to their underlying temporal deficit (Zeng et al. 2005; Zeng et al. 1999), overall scores are less, when compared to individuals with normal hearing.

The results also indicated significantly better speech identification at 10 dB SNR in presence of spectrally modulated noise with 4 ERB gaps, but not with 2 ERB gaps. At 10 dB SNR, since the effects of noise were already lesser, the additional advantage of the release from masking due to fluctuations in the masker may have resulted in improved scores for spectral modulations with 4 ERB gaps, but not with 2 ERB gaps. Even then, these individuals did not show any benefit from temporal modulations in noise at 10 dB SNR. This would imply that even at favorable noise conditions like 10 dB SNR, these individuals exhibit poorer temporal processing.

The results on amount of release from masking obtained for each of the modulated noise conditions also revealed that there is maximum release from masking with a spectrally modulated noise with 4 ERB gaps followed by the spectrally modulated noise with 2 ERB gaps at both 0 dB SNR and 10 dB SNR. There is minimal or no release from masking obtained for temporally modulated noise. The cross talk between the nerve fibers would probably have caused smearing of adjacent frequencies and hence the narrow spectral gaps (ERB 2), could have been masked relatively more by the smearing when compared to 4 ERB spectral gaps with far off frequencies. Therefore speech identification under noise with 4 ERB spectral modulations showed maximum release from masking.

Minimal or no release from masking obtained for temporal modulations in noise could also be attributed to the underlying temporal deficit caused by the dyssynchronous firing of neural impulses.

Individuals with Cochlear Hearing Loss

The individuals with cochlear hearing loss also showed maximum masking for unmodulated masker at 0 dB SNR as also reported by Leger et al. (2012). Investigators have reported that when the masker is modulated either periodically or by the speech of a single talker, speech intelligibility improves compared to when unmodulated noise is used , even if the modulated and un modulated noises have equal average powers (Festen & Plomp, 1990). The results also indicated that there was no significant difference between the speech identification under un modulated noise and spectrally modulated noise with 2 ERB gaps, i.e. these individuals could not take benefit of noise with spectral modulations with narrow 2 ERB gaps. Peters et al. (1998) reported that individuals with cochlear hearing loss perform significantly poorer in speech identification in presence of spectrally modulated noise.

This would indicate that reduced spectral resolution due to broader auditory filters in individuals with cochlear hearing loss (Glasberg & Moore, 1986) do not allow them to take benefit of narrow spectral gaps (ERB 2).

But the results also indicated that when the noise had broader spectral modulations with 4 ERB gaps, the individuals with cochlear hearing loss attained significantly better identification than under un modulated noise. It could be because the noise with 4 ERB gaps gives broader spectral gaps; to help these individuals also take benefit of release from masking. But these results of the present study are contradicting with the findings reported by Peters et al. (1998). They reported that individuals with cochlear hearing loss do not show much benefit in terms of SNR required to achieve 50% speech identification scores even when the bandwidth of spectral modulations were increased from two to four ERBNs.

Another finding of the present study was that the temporal modulations in noise showed no significant benefit when compared to steady state noise at 0 dB SNR. The reduced release from masking in presence of temporally modulated noise could be attributed to reduced frequency selectivity in such individuals (Peters et al, 1998). The broader auditory filter bandwidths in such individuals could cause reduced ability to use fine structure information and hence affect neural coding of temporal information (Glasberg & Moore, 1986) reported that a reduced release from temporal dips in noise could be due to deficit in coding the temporal fine structure cues in the signal due to poor phase locking ability of the nerve fibers. This would imply that some amount of temporal processing deficits is also exhibited by the individuals with cochlear hearing loss.

The comparison of modulated and unmodulated maskers did not show any significant differences at 10 dB SNR. This could be because, at this condition, the level of noise is inadequate to mask the high level of speech in individuals with cochlear hearing loss. At 10 dB SNR, maximum scores could be obtained even in presence of unmodulated noise condition.

The comparison of amount of release from masking obtained from various modulated maskers at 0 dB SNR also revealed that individuals with cochlear hearing loss demonstrated a significant amount of release only for the noise with 4 ERB spectral modulations. This could also be reasoned with the reduced spectral and temporal resolution in such individuals.

Individuals with Normal Hearing Sensitivity

Individuals with normal hearing sensitivity showed significantly better speech identification scores in presence of spectrally modulated noises and temporally modulated noise than when compared to the unmodulated masker. This indicates that individuals with normal hearing sensitivity have the ability to take advantage of spectral and temporal dips in the noise to understand the target speech signal (Peters et al, 1998; Festen & Plomp, 1990). The intact spectral and temporal resolution in these individuals facilitated to utilize the spectral and temporal fluctuations in the masker (Leger et al., 2012; Peters, et al., 1998).

The results also indicated that individuals with normal hearing benefited from spectral modulations with 2 ERB gaps as comparable to masker modulations of 4 ERB gaps, since the speech identification obtained under those two noise conditions were not statistically significant. This indicates that these individuals could take the advantage effectively even for narrow ERB gaps (2 ERB gaps). Thus there was no significant difference between the speech identification obtained under noises with 2 ERB and 4 ERB modulations.

When the amount of release obtained for each of the modulated masker was compared, no significant differences between the modulated noises were obtained. This implicated that individuals with normal hearing sensitivity utilized both spectral and temporal modulations in the masker to the same extend. Peters et al., (1998) and Duquesnoy, (1983) reported that normally hearing listeners can obtain very large advantage of listening in spectral and temporal dips.

Comparison of effect of various maskers across the groups

Results revealed that individuals with AD performed worst under all conditions of noise than normal hearing listeners or those with cochlear hearing loss. At 0 dB SNR, except for the spectrally modulated noise with 4 ERB gaps, under all other types of noise (modulated and un modulated), the speech identification was greatest in individuals with normal hearing sensitivity followed by those with cochlear hearing loss and then by those with AD. These results are in line with the findings reporting excessive masking effects observed in individuals with AD followed by cochlear hearing loss. Rance et al. (2007) reported that children with AD have significant perception problems in noise than when compared to peers having cochlear hearing loss.

Spectrally modulated noise with 4 ERB gaps gave the maximum speech identification scores across all three groups of individuals. This indicates that it was the easiest of all the noise conditions. Duquesnoy, (1983) reported that as the width of the spectral dips increases, the speech identification performance increases.

In the presence of spectrally modulated noise with 4 ERB gaps condition at 0dB SNR, individuals with cochlear hearing loss and those with normal hearing had similar scores. This implies that individuals with cochlear hearing loss could take the benefit of spectral modulations with 4 ERB gaps as normal hearing listeners.

Also when the amount of release from masking was compared across the 3 groups, it was found that individuals with AD differed significantly from other 2 groups. These results indicated that there was little or no amount of release obtained with temporally modulated noise in individuals with AD when compared to the other groups. Also, individuals with AD show significant release from masking in presence of spectrally modulated noise when compared to individuals with cochlear hearing loss. This also points out to the poor temporal processing in these individuals when compared to those with normal hearing (Liu et al., 2004) and cochlear hearing loss (Payton, Uchanski & Braida, 1994). A direct comparison with individuals with normal hearing cannot be made because individuals with normal hearing obtained relatively good speech identification scores even for the unmodulated noise. Hence the improvement in scores noticed with spectral modulations with 4 ERB gaps was the maximum in individuals with AD.

Conclusions

The major findings of the study indicated that individuals with Auditory dys-synchrony can extract the target speech signal when the background noise has larger spectral dips. Due to their underlying temporal processing problem they could not differentiate the temporal gaps in noise and hence perception was poorer in presence of noise. Any noise reduction strategies should incorporate large spectral dips in continuous noise to enhance the speech perception by allowing glimpses of signal. The individuals with cochlear hearing loss also performed significantly better in the presence of noise with 4 ERB spectral gaps, but the improvement noted was lesser compared to those individuals having AD. This could be attributed to the broader auditory filter bandwidths in these individuals which may have disallowed glimpses of speech when the spectral dips were narrow. It was also found that individuals with normal hearing sensitivity could utilize even smaller glimpses present in the noise. Thus, the use of a fluctuating noise in assessing speech recognition may provide us with a sensitive way of evaluating the effects of signal processing in hearing devices such as frequency-selective amplification and compression.

References

American National Standards Institute. (1996). *Specifications for Audiometers* (ANSI S3.6-1996). New York, NY.

Bernstein, J. G. W., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 125*, 3358-3372.

Buss, E., Hall III, J. W., and Grose, J. H. (2004) "Temporal fine-structure cues to speech and pure tone modulation in patients with sensorineural hearing loss," *Ear and Hearing, 25*(3), 242-50.

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons. *Journal of the Acoustical Society of America, 74*, 739-743.

Eggermont, J. (1997). Firing rate and firing synchrony distinguish dynamic from steady state sound. *Neuroreport, 8,* 2709-2713.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America, 88,* 1725-1736.

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in participants with unilateral and bilateral cochlear impairments. *Journal of the Acoustical Society of America, 79*, 1020-1033.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47,* 103-138.

Harrison, R. (1998) An animal model of auditory neuropathy. *Ear and Hearing, 19*, 355-361

Kraus, N., Bradlow, M. A., Cheatham, M. A., Cunningham, C. J., King, C. D., Koch, D. B., et al. (2000). Consequences of neural asynchrony: a case of auditory neuropathy. *Journal of the Association Research in Otolaryngology, 1*, 33-45.

Leger, A., Moore, B. C. J., & Lorenzi, C. (2012). Temporal and spectral masking release in the low- and mid-frequency range for normal-hearing and hearing-impaired listeners.*Journal of the Acoustical Society of America, 131*, 1502-1514.

Liu, S., Del Rio, E., Bradlow, A. R., & Zeng, F. G. (2004). Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America, 116*, 2374-2383.

Methi, R., Avinash, & Kumar, U. A. (2009). Development of sentence material for Quick Speech in Noise test (Quick SIN) in Kannada. *Journal of Indian speech and Hearing Association, 23*(1), 59-65.

Middelweerd, M. J., Festen, J. M., & Plomp, R. (1990). Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram. *Audiology, 29*(1), 1-7.

Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing, 11,* 133-160.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). Academic Press: London.

Moore, B. C. J., Glasberg, B. R., & Hopkins, K. (2006). Frequency discrimination of complex tones by hearing-impaired participants: Evidence for loss of ability to use temporal fine structure information. *Hearing Research, 222,* 16-27.

Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. Wiley-Interscience: Chichester, England.

Payton, K. L., Uchanski, R. M., and Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing, *Journal of the Acoustic Society of America, 95*(3),1581-1592.

Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *Journal of the Acoustical Society of America, 103,* 577-587.

Rance, G., Mc Kay, C., & Grayden D. (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing, 21*, 34-46.

Rance, G., Barker, E., Mok, M., Dowell, R., Rincon, A. & Garratt, R. (2007). Speech perception in noise for children with auditory neuropathy/dys-synchrony type hearing loss. *Ear and Hearing, 28*, 351-360.

Starr, A., Picton, T., Hood, L. J., & Berlin, C. (1996). Auditory neuropathy. *Brain, 119*, 741-753.

Starr, A., Picton, T. W., & Kim, R. (2001). Pathophysiology of auditory neuropathy. In Y. S. Sininger & A. Starr (Eds.), Auditory neuropathy (pp. 67-82). San Diego, CA: Singular.

Starr, A., Michalewski, H.J., Zeng, F.G., Brooks, S.F., Linthicum, F. Kim, et al. (2003). Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene. *Brain, 126*, 1604-1619.

Vinay, S. N. & Moore, B. C. J. (2007). Ten(HL)-test results and psychophysical tuning curves for participants with auditory neuropathy. *International Journal of Audiology, 46,* 39-46.

Wagener K., Brand T. & Kollmeier B. (2006). The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners. *International Journal of Audiology* , *45*(1), 26 - 33.

Zeng, F. G., Oba, S., Garde, S., Sininger, Y.,& Starr, A. (1999). Temporal and speech processing deficits in auditory neuropathy. *Neurology Report, 10*, 3429-3435.

Zeng, F. G., Oba, S., & Starr, A. (2001). Supra threshold processing deficits due to desynchronous neural activities in auditory neuropathy. In D.J Breebaart, A.J.M Houstma, A. Kohlrausch, et al. (Eds.), *Physiological and Psychophysical Bases of Auditory Function* (pp.365-372). Maastricht, Netherlands: Shaker Publishing BV.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J. & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology, 93*, 3050-3063.

Zeng, F. G., & Liu, S. (2006). Speech perception in individuals with auditory neuropathy. *Journal of Speech, Language and Hearing Research, 49*, 367-380.

Age Related Changes in Auditory Memory and Sequencing

[1]Mythri & [2]Asha Yathiraj

Abstract

The study aimed at investigating the effect of age and gender on auditory memory and sequencing abilities in individuals with normal peripheral hearing. The study also evaluated the relation between perceived memory problems (evaluated using a checklist) with actual memory problems (evaluated using 'Auditory Memory and Sequencing Test in Kannada'). Sixty participants, divided into three groups based on their age were recruited. Group-I and Group-II included older adults in the age ranges of 50 to 64;11 years and 65 to 80 years respectively. Group-III included younger adults in the age range of 20 to 30 years. The results of the study revealed a significant decline in memory and sequencing abilities with age. In all three groups, higher scores were obtained on the memory subtest compared to the sequencing subtest. In both the older adult groups, males performed superior to the females on the sequencing subtest. Such a gender difference was not seen in the younger group. The gender difference in the older groups was attributed to the educational difference between the males and females in these participants which was not present in the younger group. Also, in all but Group-I, the self-report scores regarding their memory abilities corresponded with their actual memory abilities.

Keywords: *Aging, auditory memory and sequencing, perceived memory problem*

Introduction

Aging has been referred to as a multidimensional process of physical, physiological and social change (Hamilton, 2006). It is reported to be deeply rooted in the genetic makeup and metabolic working of an organism (Braver & Barch, 2002) and sensitive to many environmental influences (Arking, 1991). As adults grow older, physical, sensory, emotional, psychological and social changes are reported to occur (Dugan & Kivett, 1994). The rate at which these changes occur and how they affect an individual are noted to be based on a number of factors.

Changes in the structure and function are also reported to occur throughout the peripheral and central auditory nervous system as a result of the aging process. Many investigators have examined age-related changes in processing non-speech signals (McCroskey & Kasten, 1982; Newman & Spitzer, 1983) and complex speech signals (Jerger & Hayes, 1977; Konkle, Beasley & Bess, 1977; McCroskey & Kasten, 1982; Rastatter & Hood, 1986). Karlin (1942) noted that tests of conventional auditory acuity had little value in predicting auditory behaviour in more complex social situations in older adults. The effect of degenerative changes were reported to become evident only when older listeners were perceptually stressed, such as when they were required to listen to complex signals or in a noisy environment where more complex auditory processing was required.

Normal aging has also been associated with a decline in memory abilities and the phenomenon has been termed as age-related memory impairment or age-associated memory impairment. A large number of elderly individuals have been reported to live with mild memory problems that are a part of a normal aging process (Schroder, et al., 1998). Timothy (2009) reported that the steady decline in many cognitive processes was observed across the lifespan, accelerating from the twenties or thirties. The author claimed that due to aging, attention and memory were the most affected basic cognitive functions.

Older adults have also been reported to exhibit deficits on temporal ordering tasks (Parkin, Walter & Hunkin, 1995). Neils, Newman, Hill and Weiler (1991) found that elderly individuals performed significantly poorer than the younger adults on auditory memory and sequencing of tones. Gregoire and Linden (1997) noted that a major part of the adult lifespan was characterized by slight decline in memory abilities. Studies by Mitrushina and Satz (1991), Youngjohn and Crook (1993), Small, Stern, Tang and Mayeux (1999), Oberauer, Wendland and Kliegl (2003), and Moral, Tomas, Bataller, Oliver and Navarro (2010) have also documented a significant decline in auditory memory with age.

Several reasons have been speculated by Light and Leah (1999) to explain why older adults use less effective encoding and retrieval strategies as they age. The first was the 'disuse' view, which implied that memory strategies were used less by older adults as they moved further away from the educational system. Second was the 'diminished attentional capacity' hypothesis, which meant that older people engaged less in self-initiated encoding due to reduced attentional capacity. The third reason was the 'memory self-efficacy', which indicated that older people did not have confidence in their own memory performances, leading to poor consequences.

The ability to encode new memories of events or facts

[1]Email: my3hm.66@gmail.com,
[2]Professor of Audiology, Email: asha_yathiraj@redffmail.com

has been shown to decline in both cross-sectional and longitudinal studies (Park, 1996; Park 2002; Hedden & Gabrieli, 2004). Chisolm, Willott and Lister (2003) reported that the most prevalent but most often overlooked skill deficiency in elderly subjects was auditory memory. In a study by Schroder et al. (1998), the prevalence of age-associated memory impairment was found to be 13.5% in individuals between the age range of 60 to 64 years. Hanninen et al. (1996) reported a higher (38.4%) prevalence of age-associated memory impairment using the National Institute of Mental Health criteria in an elderly population in Finland.

The vast majority of the studies regarding the auditory memory problems of the elderly have been done in other countries. In India, the auditory memory and sequencing problems have been studied more in children. The tests for auditory memory and sequence, developed in India (Yathiraj & Vijayalakshmi, 2006; Yathiraj & Mascarenhas, 2003) have focused on children and not on the older generation.

A high correlation between memory skills and educational level has been found (Gathercole, Pickering, Knight & Stegmann, 2004). It was found in the National Sample Survey report (2004-2005) that only 63.6% of the total population of India is literate. The literacy percentage of males and females was found to be 78.0% and 51.10% respectively. Hence, there is a need to see if age related changes seen in the older population in India are similar to that found in the other parts of the world where the literacy level is noted to be higher.

Additionally, there is a need to have information regarding deficits in auditory memory and sequencing since these aspects can affect the audiologic rehabilitation goals and outcomes for older adults. Knowledge about the individual's auditory memory and sequencing skills can help to plan the rehabilitation goals or modify the activities/tasks appropriately. Hence, it is necessary to see the relation between age and auditory memory.

Rehabilitation outcome can also be affected if there is a mismatch between the perceived degrees of memory impairment, as reported by the client and the actual degree of impairment. Therefore, it is essential to study if there exist any relation between perceived and the actual degree of memory impairment in older individuals.

Thus, the present study aimed to investigate the effect of age and gender on auditory memory and sequencing in older adults with normal hearing sensitivity and to compare them with young adults. The study also aimed to examine the relation between the perceived degree of memory impairment and the actual degree of impairment.

Method

Participants

The participants were divided into 3 groups based on their age (Group I, Group II and Group III). Each group included 20 participants with 10 males and 10 females. Group-I and Group-II included older adults in the age ranges of 50 to 64;11 years and 65 to 80 years respectively. Group-III included normal hearing individuals in the age range of 20 to 30 years. The educational levels of all the participants were noted and are presented in the Table 1.

The participants were native speakers of Kannada. They had normal AC and BC pure-tone thresholds after applying a correction factor for age as recommended by Indrani (1981), whenever required. Their speech identification score was 80% or more on the 'Phonemically

Table 1: Educational levels of the participants

Educational level	Group I (50 to 64;11 years)		Group II (65 to 80 years)		Group III (20 to 30 years)		TOTAL
	Males	Females	Males	Females	Males	Females	
Primary	1	0	1	2	0	0	4
Secondary	1	4	2	5	0	0	12
High school	1	2	2	3	0	0	8
Pre-university	2	4	0	0	0	0	6
Graduation	4	0	5	0	6	5	20
Post-graduation	1	0	0	0	4	5	10
TOTAL	10	10	10	10	10	10	60

balanced Kannada word test' developed by Yathiraj and Vijayalakshmi (2005). None of the participants reported of any history of middle ear pathology or any major neurological problem. An informed consent was taken from all the participants prior to carrying out the evaluations.

Material

To obtain information about early signs of dementia, a checklist was developed based on the information reported in literature and the opinion of experienced Speech and Hearing professionals. The initial checklist contained 12 questions which required responses on a 3-point scale ('never', 'sometimes' and 'always'). A symptom that occurred less than 25% of the time was required to be labelled 'never', while those that occurred 25% to 75% and more than 75% of the time were required to be labelled 'sometimes' and 'always' respectively. Further, the checklist was scored by awarding a response 'never' a score of 0, while 'sometimes' was scored 1 and 'always' was scored 2. Thus, the total possible score ranged between 0 and 18 on the developed 'Memory ability checklist' (Appendix).

Item validity was checked by obtaining the opinion of five speech and hearing professionals who had at least 10 years of experience in the area of cognition. After incorporating the modifications and suggestions of the speech and hearing professionals, the checklist had nine questions. As no changes were recommended for the procedure to obtain the responses, the 3-point rating scale was retained. To test the auditory memory and sequencing abilities of the participants, they were tested using the 'Auditory Memory and Sequencing Test in Kannada' developed by Yathiraj and Vijayalakshmi (2006). The test contained four lists of words with different inter-stimulus intervals (250 msec, 500 msec, 750 msec and 1 sec). Each list commences with a three-word token and gradually increased to an eight-word token with a total of twenty tokens. The list with an inter-stimulus interval of 500 msec was used for the present study.

Instrumentation

Madsen Orbiter-922 type-I diagnostic audiometer with calibrated TDH-39 headphones was used to estimate the air conduction thresholds and to carry out speech audiometry (ANSI S3.6, 1996). Calibrated Radio Ear B-71 bone vibrator was used to estimate bone conduction thresholds. The same audiometer was also used to route the stimuli for the auditory memory and sequencing test in Kannada from a laptop to loud speakers. A calibrated Grason StadlerInc-Tympstar, clinical immittance meter used to rule out any middle ear pathology.

Test Environment

All tests were administered in an acoustically treated suite. It was ensured that the noise levels were within the permissible limits as recommended by ANSI S3.1 (1991).

Procedure for Participant Selection

Pure-tone thresholds were obtained in octave intervals between 250 Hz to 8000 Hz for air conduction and between 250 Hz and 4000 Hz for bone conduction using the modified Hughson - Westlake procedure (Carhart & Jerger, 1959). Speech identification scores (SIS) were obtained under headphones using the phonemically balanced word list (Yathiraj & Vijayalakshmi, 2005) at 40 dB SL. Participants with a pure-tone threshold of less than 25 dB HL, after applying a correction factor for age as recommended by Indrani (1981) and a SIS of greater than 80% were selected.

Tympanometry and reflexometry were carried out to rule out any possibility of middle ear pathology using a 226 Hz probe tone. Ipsilateral and contralateral reflexes at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz were obtained. Participants with type-A tympanogram with bilateral reflexes present were considered for further assessment.

The developed 'Memory ability checklist' was administered to obtain demographic details of the participants and to obtain information regarding the perception of their memory abilities. Each participant was required to respond on the three point scale. Only those participants with a total score of less than 9 were selected. This was done to avoid including those with considerable memory problems who may not be able to give adequate responses on the Auditory Memory and Sequencing Test.

Procedure for administering the Kannada Auditory Memory and Sequencing Test

The Kannada Auditory Memory and Sequencing Test (Yathiraj & Vijayalakshmi, 2006) was administered on participants who met the participant selection criteria. A CD containing the test stimuli were played on a laptop with Intel Core i3 processor. The signal from the laptop was fed to the CD input of the Madsen Orbiter-922 type-I diagnostic audiometer. The output of the audiometer was given to a loud speaker which was placed 1 meter from the head of the participants at 0o azimuth. The signal was presented at 40 dB HL.

The participants were instructed to listen to the group of words present in each token and repeat them in the same order. A score of one was awarded for every correct word that was recalled. An additional score of one was awarded if the words recalled were in the correct sequence. The responses were noted on a scoring sheet and the total score for the memory subtest and the sequencing subtest was calculated.

Analysis

The raw scores obtained from the 60 participants on the 'Kannada auditory memory and sequencing test' and the 'Memory ability checklist' were tabulated. The data thus obtained was subjected to statistical analyses, using SPSS (Version 18). MANOVA was done to see the effect of age and gender on the scores of the auditory memory and sequencing subtests. To study the effect of age and gender on the total score of the 'Kannada auditory memory and sequencing test', ANOVA was carried out. Non-parametric Kruskal-Wallis test was used to see the impact of age on the scores obtained on the 'Memory ability checklist'.

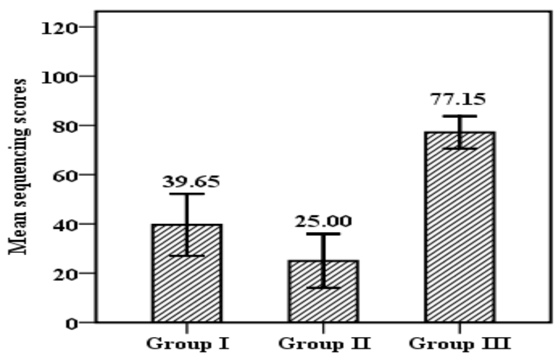
Results

Based on the statistical analyses, a comparison was made between the scores obtained across the three age groups (50 to 64;11 years, 65 to 80 years & 20 to 30 years) for four different scores. The scores included the auditory memory subtest score, auditory sequencing subtest score, total auditory memory and sequencing score, and the 'Memory ability checklist' score. A comparison was also made between the genders on the same parameters.

Effect of Age

Effect of age on the auditory memory subtest score: The mean and standard deviation of the auditory memory subtest score was determined for each age group (Figure 1). From the figure it is evident that the mean memory score of the two groups of older adults was lesser than that of the younger adults.

To check if this difference was statistically significant, MANOVA was carried out. The results indicated that there was a significant main effect [$F(2, 54) = 216.113$, $p < 0.05$]. Further, Duncan post-hoc test was used to see whether each age group differed significantly from



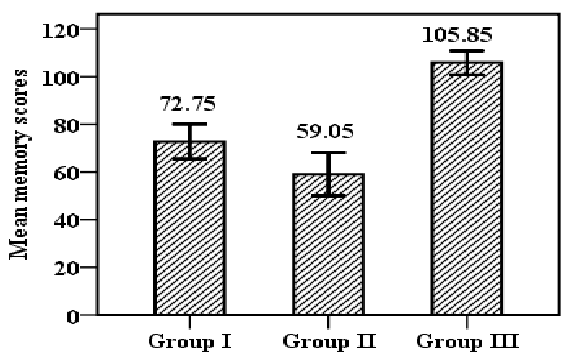
*Note:# Maximum score is 118; * $p < 0.05$*

Figure 2: Mean and standard deviation of the sequencing subtest scores for the three age groups.

the other on the memory subtest scores. The results indicated that the scores of Group-I and Group-II differed significantly from the Group-III ($p < 0.05$) and differed significantly from each other ($p < 0.05$).

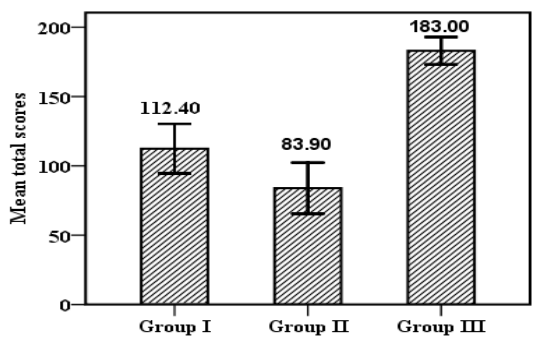
Effect of age on the auditory sequencing subtest score: From the descriptive statistics (Figure 2) it is apparent that the mean sequencing score of the younger adults (Group III) was higher than that of the two older groups (Group I & Group II). Also, the variability for the older adults was more compared to the younger adults. The results of MANOVA also indicated a significant main effect of group [$F(2, 54) = 154.082$, $p < 0.05$]. The findings of the Duncan post-hoc test showed that the mean sequencing scores of Group-I and Group-II differed significantly ($p < 0.05$) from the Group-III. Additionally, the two older groups also differed significantly from each other ($p < 0.05$).

Effect of age on the total auditory memory and sequencing score: The findings of the descriptive statistics (mean and SD) for the total memory and sequencing score across age groups are presented in the Figure 3. It can be observed that the scores decreased with in-



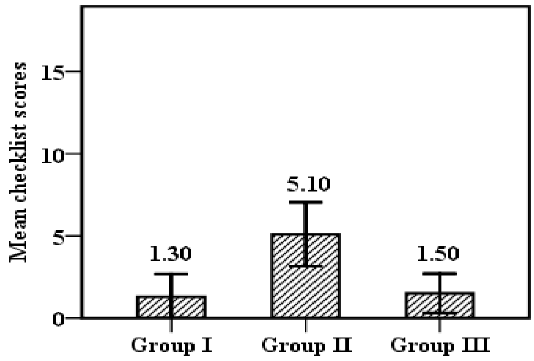
*Notes: # Maximum score is 118; * $p < 0.05$*

Figure 1: Mean and standard deviation of the memory subtest scores for the three age groups.



*Note: # Maximum score is 236; * $p < 0.05$*

Figure 3: Mean and standard deviations of the total auditory memory and sequencing scores for the three age groups.



*Note: # Maximum score is 18; * $p < 0.05$*

Figure 4: Mean and standard deviations of the scores on the 'Memory ability checklist' for the three age groups.

crease in age. To determine the effect of age on these total scores of the 'Kannada auditory memory and sequencing test' ANOVA was carried out. A statistically significant main effect for groups [$F(2, 54) = 224.600$, $p < 0.05$] was seen. Duncan post-hoc test indicated a significant difference in the mean total auditory memory and sequencing scores of Groups I and II ($p < 0.05$), Groups I and III ($p < 0.05$) as well as between Groups II and III ($p < 0.05$).

Effect of age on the 'Memory ability checklist' score: The mean and SD of the 'Memory ability checklist' scores of the three age groups are depicted in Figure 4. To compare the age effect on the 'Memory ability checklist' score, non-parametric Kruskal-Wallis test was carried out. The results showed a significant age effect [$\chi^2(2) = 31.344$, $p < 0.05$]. The findings of the post-hoc analysis carried out using the non-parametric Mann-Witney test showed a significant difference between Group-I and -II ($p < 0.05$) and Group-II and -III ($p < 0.05$) but not between Group-I and -III ($p > 0.05$).

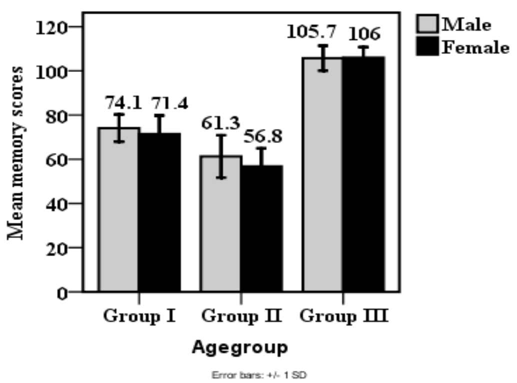
Effect of Gender

Effect of gender on the auditory memory subtest score: The mean and standard deviation of the auditory memory subtest score was determined across genders for each of the three age groups (Figure 5). The MANOVA results indicated that there was no significant gender effect [$F(1, 54) = 1.481$, $p > 0.05$] when the three age groups were combined. To determine if this lack of gender difference was maintained for each of the three age groups, group wise comparison using MANOVA was done. It was seen that in each of the age groups, both males and females performed equally ($p > 0.05$) on the memory subtest.

Effect of gender on the auditory sequencing subtest score: The descriptive statistics of the sequencing subtest (Figure 6) indicated that the females in the older two groups performed poorer than the males. Such a difference was not observed in the younger age group. The results of MANOVA indicated a significant difference [$F(1, 54) = 6.680$, $p < 0.05$] between the genders, when the age groups were combined. However, the group-wise comparison revealed that the significant difference between genders was present in Group-I ($p < 0.05$) and Group-II ($p < 0.05$) and not in Group-III ($p > 0.05$).

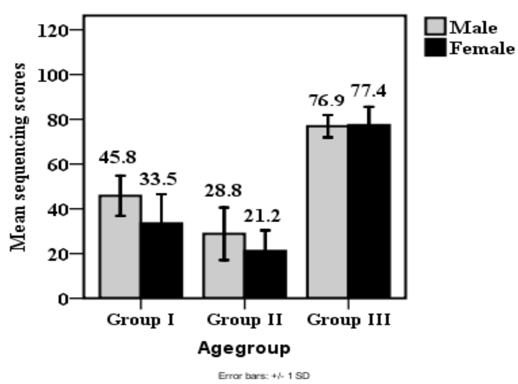
Effect of gender on the total score: ANOVA was carried out to see the effect of gender on the total score of the 'Kannada auditory memory and sequencing test'. The overall results indicated significant difference [$F(1, 54) = 4.861$, $p < 0.05$] between genders. However, a significant difference was present only in the older two adult groups ($p < 0.05$) and not in younger adults group ($p > 0.05$) on the Duncan post-hoc test (Figure 7).

Effect of gender on the 'Memory ability checklist' score: Figure 8 shows the mean 'Memory ability checklist' scores across genders for all the three age groups. To study the gender effect on the 'Memory ability check-



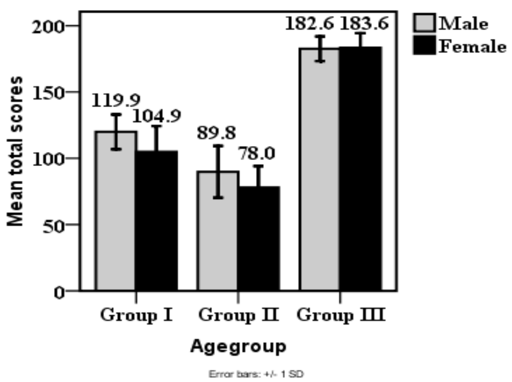
Note: # Maximum score is 118

Figure 5: Mean memory subtest scores for males and females across the three age groups.



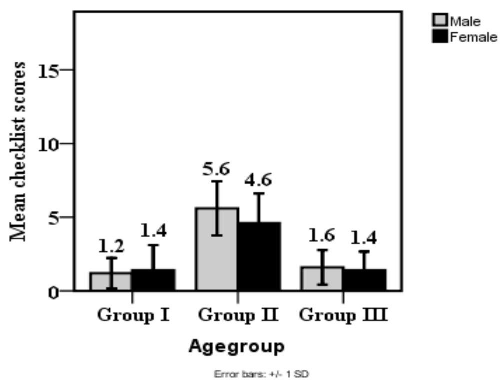
*Note: # Maximum score is 118; * $p < 0.05$*

Figure 6: Mean sequencing subtest scores for males and females for the three age groups.



*Note: # Maximum score is 236; * $p < 0.05$*

Figure 7: Mean total memory and sequencing scores for males and females across the three age groups.



Note: # Maximum score is 18

Figure 8: Mean scores on the 'Memory ability checklist' for males and females across the three age groups.

list' score, non-parametric Kruskal-Wallis test was carried out. The results showed no significant difference [χ^2 (2) = -0.158, p>0.05] between genders across all the three age groups.

The findings of the study are discussed in the light of the available literature. This is done to see if the findings support or refute the findings of studies reported in literature.

Discussion

The comparison of the scores obtained on the auditory memory and sequencing test and the 'Memory ability checklist' are discussed for three age groups as well as across the genders. The comparison of the 'Memory ability checklist' scores with the total auditory memory and sequencing scores for each age group are also discussed.

Comparisons of memory and sequencing abilities across age groups

The results of the present study showed that the younger adults aged 20-30 years (Group-III) performed better than the two older groups aged 50 to 64;11 and 65 to 80 years (Group-I & -II respectively). This was observed on auditory memory subtest as well as the sequencing subtest. Similar results are also seen on the total auditory memory and sequencing scores.

These results are in agreement with those of Anders, Fozard and Lillyquist (1972) and Neils et al. (1991). They reported of considerable decrement in the memory and sequencing scores of the elderly compared when compared younger adults on a recognition task and free recall task respectively. Kester, Benjamin, Castel and Craik (2002) also opined that the elderly experience trouble retrieving information from memory, particularly when retrieval required effortful processing, as in un-cued recall.

In the current study, a comparison of the two older groups of adults (Groups I & II) revealed that the younger of these two groups (50 to 64;11) got significantly higher scores than the older of these groups (65 to 80 years). This was observed on the memory as well as sequencing subtests.

Similar findings have been reported by Park (1996) in a longitudinal comparison study. The study revealed that age-related changes from age 20 to 60 tended to be small. In contrast, changes after the age of 60 had a steeper slope. Park (2002) too found that memory scores showed a linear life-long decline with an accelerated decline in the later decades. Several other studies have also reported of a generalized slowing in brain function which resulted in a decline in problem solving, reasoning, memory, and language (Cerella, 1990; Lindenberger & Baltes, 1994; Salthouse, 1996). Additionally, a slowing of behaviour in old age has also been documented (Birren, 1965; Salthouse, 1991, 1996).

Age related decline in memory and sequencing has been ascribed to several reasons in literature. Salthouse (1996) suggested that age-related impairments on tasks that do not have an obvious speed component, such as free recall tasks, could be explained via the simultaneity mechanism where the products of earlier processes were lost before later processes were carried out. The reduction in the memory and sequencing scores with aging has also been attributed to inadequate signal processing due to aging of the sensory systems by Murphy, Craik, Li and Schneider (2000). Yet another reason for a reduction in performance with aging, according to Craik and Byrd (1982), had been the depletion of attentional resources available for cognitive processing. Further, Kester et al. (2002) noted that age-related decrement in executive control over cognitive processes has lead to a decline in memory ability in the elderly.

Thus, the findings of the present study are in consonance with that reported in literature. These results add to the corpus of findings regarding age related decline in auditory memory and sequencing. The current study highlights that such age related changes are universal and not restricted to certain regions or communities. Further, it can be construed that the age related changes seen in memory and sequencing performance in the present study could be due to a combination of inadequate signal processing, attention reduction and cognitive decrement. Additionally, in agreement with Park (2002), these changes continue to decline with advancement in age.

Comparisons between memory and sequencing scores across age groups

Overall, in the present study, all three groups obtained higher scores on the memory sub-test compared to the sequencing sub-test, indicating that the latter task was a more challenging one. The drop in score was similar for the two older groups (33.1 & 34.05 for Groups I & II respectively). However, their drop in score was more than that of seen in the younger group (28.7). This indicates that the older two groups found the sequencing task more difficult than the younger group.

Similar findings have been reported by Yathiraj and Vijaylakshmi (2006) in children aged 11 to 12 years. The authors found that the memory subtest scores ranged between 101 to 105, with the maximum attainable score being 118. However, the scores for the sequencing subtest dropped to 69 to 78 for the same maximum attainable score of 118.

Comparison across gender in each age group

It was found in the present study that both males and females performed equally on the memory subtest. Similar findings were seen across all three age groups. This finding is consistent with that of Susan, Susan, Benjamin and Hannah (2004) who found no significant between genders on a working memory task.

However, on the sequencing subtest, the present study revealed that the performance of the female participants was poorer than that of the male participants. This was observed in the two older adult groups (Group-I & -II) but was absent in the younger adult group (Group-III) where both genders performed equally. A possible reason for the presence of a gender difference in the older two groups could be due to the educational differences in the males and females. From the Table 1 presented in the method section, it is evident the males in both the elderly groups had higher educational levels when compared to the females of the same age group. Such a difference in educational level was not present between the genders in the younger adult group.

The effect of education on the performance of males and females has been reported by Coffey, Saxton, Ratcliff, Bryan, and Lucke (1999). The authors found that each year of education was associated with an increase in peripheral CSF volume (a marker of cortical atrophy) of 1.77 mL ($p < 0.03$) in a nonclinical population. In the present study, since the females had comparatively less education than the males, they were likely to have a greater degree of cortical atrophy which in turn could have resulted in poor performance on the sequencing subtest.

In contrast to the findings of the present study, Alexander, Packard and Peterson (2002) and Lowe, Mayfield and Reynolds (2003) documented better scores in young and older females on various memory and sequencing tasks. Subject variability, material used for the task and education level of the subjects may have accounted for the difference in the findings between their study and that of the present one.

Craik and Byrd (1982) reported that difficult tasks required more attentional capacity than simpler tasks. In the current study, the sequencing task was more taxing. This is evident from the lower scores on this task compared to the memory subtest. Since the sequencing subtask was relatively more complex than the memory subtask, the differences in the gender could probably be picked-up with the former and not with the latter subtest.

Comparisons of the total auditory memory and sequencing test scores and 'Memory ability checklist' score

The results of the study revealed that all the three age groups differed significantly from each other on the total auditory memory and sequencing test score. The younger adults (20 to 30 years) performed better than the two groups of older adults. Among the two groups of older adults, Group-I (50 to 64;11 years) performed superior to the Group-II (65 to 80 years).

The younger adults (Group-III) had lowest scores on the 'Memory ability checklist' which indicated that they did not perceive themselves as having any memory problems. This was in agreement with their scores obtained on the memory and sequencing test. The 65 to 80 years old participants (Group-II) had the highest scores on the 'Memory ability checklist' which was also evident from their poor scores on the Kannada auditory memory and sequencing test. On the other hand, the participants in the Group-I, aged 50 to 64;11 years, did not report of any decline in their memory ability. However, their test scores were significantly low when compared to the younger adults.

The above findings are in accordance with that of Taylor, Miller and Tinklenberg (1992). The authors found that among the older adults (< 60 years), the decline in memory ability is not significant enough to cause an increase in the scores of self-report questionnaires. Hertzog and Dixon (1994) reported that the self-reports of the older individuals do not necessarily correspond with their actual memory ability unless the memory problem start occurring frequently. Hence, it can be concluded that a memory test is able to detect a decline in memory and sequencing abilities before the individual starts perceiving the deficit.

Conclusions

Overall, from the results of the present study it can be inferred that memory and sequencing abilities show a linear life-long decline with an accelerated decline in the later decades and these age related changes are universal and not restricted to certain regions or communities. In addition, the educational level of the participant plays an important role on the performance of complex tasks such as sequencing. In elderly individuals aged less than 65 years, self-report scores of the memory abilities may not necessarily correspond with their actual memory abilities.

From the present study it is clear that memory abilities deteriorate with age. This must be kept in mind while counselling / rehabilitating older adults Also, due to the possible influence of education in the performance of complex memory tasks, emphasis needs to be given to encourage higher levels of educations in all individuals.

References

Alexander, G. M., Packard, M. G., & Peterson, B. S. (2002). Sex and spatial position effects on object location memory following intentional learning on object identification. *Neuropsychologia, 40*, 1516-1522.

American National Standards Institute (1991). *American National Standard Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms*. ANSI S3.1. New York: American National Standards Institute, Inc.

American National Standards Institute (1996). *American National Standard: Specification for Audiometers*, ANSI S3.6. New York: American National Standards Institute, Inc.

Anders, T., Fozard, J., & Lillyquist, T. (1972). Effects of age upon retrieval from short-term memory. *Developmental Psychology, 6,* 214-217.

Arking, R. (1991). *Biology of aging: Observations and principles.* Englewood Cliffs, NJ: Prentice-Hall.

Birren, J. E. (1965). Age changes in speed of behavior: its central and neurophysiological correlates. In Welford, A.T. & Birren, J. E. (Eds), *Behavioural Aging and the Nervous System* (pp.191-216). Springfield, IL: Charles C. Thomas.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biological behaviors, 26,* 809-817.

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

Cerella, J. (1990). Aging and information-processing rate. In Birren, J. E. & Schaie, K. W. (Eds), *Handbook of the psychology of aging,* (pp.201-221). San Diego, CA: Academic Press.

Chisolm, T. H., Willot, J. F., & Lister, J. J. (2003). The aging auditory system: Anatomic and physiologic changes and their implication for rehabilitation. *International Journal of Audiology, 42*, 2S3-2S11.

Coffey, C. E., Saxton, J. A., Ratcliff, G., Bryan, R. N., & Lucke, J. F. (1999). Relation of education to brain size in normal aging: implications for the reserve hypothesis. *Neurology*, *53*, 189-96.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: the role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.) *Aging and Cognitive Processes,* Hillsdale, NJ: Erlbaum, 191-211.

Dugan, E., & Kivett, V. R. (1994). The importance of emotional and social isolation to loneliness among very old rural adults. *The Gerontologist, 34*, 340-346.

Gathercole, S. E., Pickering, S. J., Amridge, B. & Wearing, H. (2004). The structure of working memory from 4 to 15years of age. *Developmental Psychology, 40*, 177-190.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18,* 1-16.

Gregoire, J., & Linden, M. V. (1997). Effects of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4*, 140-149.

Hamilton, S. I. (2006). *The Psychology of Ageing: An Introduction*. London: Jessica Kingsley Publishers.

Hanninen, T., Koivisto, K., Reinikainen, K. J., Helkala, E. L., Soininen, H., Mykkanen, L., Laakso, M., & Riekkinen, P. J. (1996). Prevalence of ageing-associated cognitive decline in an elderly population. *Age and Ageing, 25*, 201-205.

Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews Neuroscience, 5*, 87-

96.

Hertzog, C., & Dixon, R. A. (1994). Metacognitive development in adulthood and old age. In J. Metcalfe & Shimamura (Eds.), *Metacognition: Knowing about Knowing*(pp.227-251). Cambridge, MA: MIT Press.

Indrani. (1981). *Hearing by air conduction as a function of age and sex in Indians*. Unpublished Master's Dissertation, University of Mysore. India.

Jerger, J., & Hayes, D. (1977). Diagnostic speech audiometry. *Archives of Otolaryngology, 103,* 216-222.

Karlin, J. E. (1942). A factorial study of auditory function. *Psychometrika, 7*, 251-279.

Kester, J. D., Benjamin, A. S., Castel, A. D. & Craik, F. I. M. (2002). Memory in elderly population. In Baddeley, A., Wilson, B. & Kopelman, M. (Eds.), *Handbook of hearing disorders* (2[nd] (Ed.) (pp.543-568). London: Wiley.

Konkle, D. F., Beasley, D. S. & Bess, F. H. (1977). Intelligibility of time-altered speech in relation to chronological aging. *Journal of Speech and Hearing Research, 20,* 108-115.

Light, C., & Leah, P. A. (1999). Memory and Aging: Four Hypotheses in Search of Data. *Annual Reviews of Psychology.* Annual Reviews Inc. Retrieved from http://www.annualreviews.org/doi/pdf/10.1146/annurev.ps.42.020191.002001 on 9 May 2011.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*, 339-355.

Lowe, P. A., Mayfield, J. W., & Reynolds, C. R. (2003). Gender differences in memory test performance among children and adolescents. *Archives of clinical neuropsychology, 18*, 865-878.

McCroskey, R. L., & Kasten, R. N. (1982). Temporal factors and the aging auditory system. *Ear and Hearing, 3*, 124-127.

Mitrushina, M., & Satz, P. (1991). Changes in cognitive functioning associated with normal aging. *Archives of Clincal Neuropsychoylogy, 6*, 49-60.

Moral, J. C., Tomas, J. M., Bataller, S., Oliver,A., & Navarro, E. (2010). Comparison between Spanish young and elderly people evaluated using Rivermead Behavioural Memory Test. *Aging Neuropsychology and Cognition, 17*, 545-55.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging, 15,* 323-334.

National Sample Survey, (2004-2005). *61[st] Round Survey Report.* Ministry of Human Resource Development. Government of India.

Neils, J., Newman, C. W., Hill, M., & Weiler, E. (1991). The Effects of Rate, Sequencing, and Memory on Auditory Processing in the Elderly. *Journal of Gerontology: Psychological Sciences, 46*, 71-75.

Newman, C. W., & Spitzer, J. B. (1983). Prolonged auditory processing in the elderly: Evidence from the backward recognition-masking paradigm. *Audiology, 22*, 241-252.

Oberauer, K., Wendland, M., & Kliegl, R. (2003). Age difference age differences in working memory-The roles of storage and selective access. *Memory & Cognition, 31*(4), 563-569.

Park, D. C. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging, 11,* 621-637.

Park, D. C. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*, 299-320.

Parkin, A., Walter, B., & Hunkin, N. (1995). Relationships between normal aging, frontal lobe function, and memory for temporal and spatial information. *Neuropsychology, 9,* 304-312.

Rastatter, M. P., & Hood, S. B. (1986). Simple motor and phonemic processing reaction times of elderly subjects. *The Journal of Auditory Research, 26,* 157-166.

Salthouse, T. A. (1991). *Theoretical Perspectives on Cognitive Aging.* Hillsdale, NJ:Erlbaum.

Salthouse, T. A. (1996). General and specific speed mediation of adult age differences in memory. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, SIB,* 30-42.

Schroder, J., Kratz, B., Pantel, J., Minnemann, E., Lehr, U., & Sauer, H. (1998). Prevalence of mild cognitive impairment in an elderly community sample. *Journal of Neural Transmission, 4,* 51-59.

Small, S. A., Stern.Y., Tang, M., & Mayeux, R. (1999). Selective decline in memory function among healthy elderly. *Neurology, 52,* 1392-1396.

Taylor, J. L., Miller, T. P. & Tinklenberg, J. R. (1992). Correlates of memory decline: a 4-year longitudinal study of older adults with memory complaints. *Psychology and Aging, 7,* 185-193.

Timothy, S. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*, 507-514.

Yathiraj, A., & Mascarenhas, K. E. (2003). *'Auditory Sequencing Test'.* Material developed at the Department of Audiology, AIISH, Mysore.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *'Phonemically balanced word list in Kannada'.* Unpublished Master's Dissertation, University of Mysore. India.

Yathiraj, A., & Vijayalakshmi, C. S. (2006). *'Kannada Auditory Memory and Sequencing Test'.* Material developed at the Department of Audiology, AIISH, Mysore.

Youngjohn, J. R., & Crook, T. H. (1993). Learning, forgetting, and retrieval of everyday material across the adult life span. *Journal of Clinical and Experimental Neuropsychology, 15,* 447-460.

APPENDIX

MEMORY ABILITY CHECKLIST

Name: Age:
Gender:
Date of birth:
Education:
Occupation:
Home address & telephone no:

Sl. no.	Questions	**Usually** (Greater than 75%of times)	**Sometimes** (Greater than 25% but less than 75%)	**Never** (Less than 25% of the time)
1.	While talking to people, do you have difficulty finding the right word?			
2.	Do you have difficulty in remembering important dates?			
3.	Do you have difficulty in remembering names of familiar objects?			
4.	Do you often forget why you visited a particular place?			
5.	Do you have difficulty in recalling digits, especially your phone/ vehicle/ door number?			
6.	Do you misplace things very often?			
7.	Do you lose your way while going for a walk or driving in your neighbourhood?			
8.	Have your family members or friends told you that you are repeating the same things over and over again?			
9.	Do you have difficulty recalling the correct order of information you just heard?			

Developmental Changes of Audio-Visual Integration: A Cross Sectional Study

[1]Preeti Sahu & [2]Geetha C.

Abstract

The study aimed to evaluate the developmental changes of audio-visual integration using McGurk phenomenon and the integration of voicing cue in normal hearing Kannada speaking participants. One hundred and fifty participants, age ranging from 6 years to 20 years were chosen for the study and were divided into six groups (25 participants in each group) based on age. Auditory and visual recording of six stop consonants in combination with vowel /a/ was done and stimuli were prepared for four conditions: audio only (AO); visual only (VO); audio-visual congruent (AV+); and audio-visual incongruent- for studying McGurk effect and audio-visual integration of voicing cue. Two way MANOVA and test of equality were used for analysis of data. Results revealed that all groups performed equally in AO condition. In VO condition, better responses were observed in 6 to7.11 year olds and 16 to 20 year olds. Audio-visual congruent condition resulted in 100% correct response by all the age groups. In McGurk task, 8 to 9.11 year olds performed poorly. There was no significant difference between the performance of 6 to 7.11 year old children and of 9 years or older. It can be concluded that the developmental changes of audio-visual integration of Indian children are different from the developmental changes reported in western studies. The assessment of phonology, reading and writing skills of the children whilst looking into the audio-visual integration will give better insight about the process. It can also be concluded that it is important to study the female and male participants separately and, also, to have a task which is more demanding such as testing in the presence of noise.

Keywords: *McGurk effect, Audio-visual integration, Voicing, Congruent, Incongruent.*

Introduction

Speech perception is an audio-visual phenomenon, in both degraded and no degraded conditions. The influence of vision on speech perception is evident by McGurk effect. This effect demonstrates the occurrence of sensory integration by presenting conflicting auditory and visual information, for example, an auditory /ba/ dubbed onto the articulatory movement for /ga/, results in the perception of /da/. This suggests that auditory-visual speech is perceived as a 'whole' perceptual unit rather than as a separate unimodal feature (Green & Kuhl, 1991). There are different factors that could affect the process of audiovisual integration. One of the major factors is age. McGurk and MacDonald (1976), in their earlier studies, noted that significantly fewer children, than adults show an influence of visual speech on perception. Further, research on developmental changes of speech perception, from infancy to young adulthood, has shown a general trend of increasing use of visual information (Massaro, 1984) and increasing audiovisual integration over a period of time (McGurk & Macdonald, 1976; Rosenblum, Schmuckler & Johnson, 1997; Wightman, Kistler, & Brungart, 2006).

This pattern of result has been replicated in studies done by different researchers (Massaro, 1984; Massaro, Thompson, Barron & Laren, 1986; Sekiyama & Burnham, 2004; Desjardins, Rogers & Werker, 1997; Dupont, Aubin & Menard, 2005; Hockley & Polka, 1994; rightman et al., 2006). The reason for this may be a result of ongoing peripheral vision development (Massaro, 1984). Massaro et al. (1986) also found that children are poorer lip readers than adults, which results in lesser visual influence rather than attention playing a role.

The age by which visual influence completes has also been researched upon. Few studies have observed benefit from visual speech by the pre-teen/teenage years (Conrad, 1977; Dodd, 1977, 1980; Hockley & Polka, 1994), with one report citing an earlier age of 8 years (Sekiyama & Burnham, 2004). Whereas, Iyer (2005), found that younger children (7 year olds female & 11 years old male listeners) were more likely to use auditory cues for speech perception, while older children (11 years old female & 15 years old male listeners) used both auditory and visual information in poor listening conditions. There is another study which has compared adults with children till the age of 11 years, assuming that there will not be much changes in the auditory integration process above 11 years of age (Sekiyama & Burnham, 2004).

However, there could be changes in audiovisual integration even after 11 years of age which also may vary depending on the gender. Result obtained by Iyer's study is in support of this view. From the above, it can be observed that, there are equivocal results seen with respect to age by which children achieve adult-like performance. Hence, studying audiovisual integration across age group, that is, in children even above 11 years of

[1]Email: preetisahu50@gmail.com,
[2]Lecturer in Audiology, Email: geethamysore.cs@gmail.com

age will help in better understanding of developmental changes of audiovisual integration process in normal hearing participants. This can also be useful to evaluate and rehabilitate different pathological conditions.

The auditory visual integration as demonstrated by McGurk effect was also found to be dependent on the linguistic experience of the person. It has been shown that Japanese speakers, more frequently notice the incompatibility between auditory and visual cues than the English speakers (Sekiyama & Burnham, 2004). The reason for this is that Japanese language has simple phonological structure and hence, promotes better auditory only speech perception early in development. Thus, while Japanese children may experience audiovisual interaction to some extent, they may not integrate visual information because the phonological environment does not demand it. Hence, there clearly, is an influence of the language on the integration ability.

Further, it has been found that preschool children who show substitution errors in articulation are less influenced by visual cues than those children who can produce consonants correctly during verbal communication (Desjardins et al., 1997). Children with articulatory errors were poorer at speech reading, and had a lower degree of visual influence in auditory visual speech perception, when compared to the non-substituting children. Because the integration of auditory and visual stimuli depends on the linguistic and cross linguistic experience (Sekiyama & Burnham, 2004), it is clear that articulatory experience affects speech reading and the degree of visual influence in auditory visual speech perception.

However, the studies evaluating the audio-visual integration have mainly investigated Western and Japanese population. The phonological system of Indian languages and the articulatory experience of Indian population are very different from English and Japanese languages. The conclusions drawn from those studies may not hold good to the Indian population, as the phonological structure of Indian languages are very different.

In the Indian context, a research study perceptually evaluated the auditory integration of ten Hindi speaking normal individuals in the age range of 18 to 25 years in CVC context (Sarika, 1995). The results of this study revealed that the auditory and visual mode influence each other. Further, when the stimulus was in the back of the oral cavity, the subject were likely to perceive the in between stimulus, i.e., neither auditory nor visual stimulus. As it can be noted, in the above study, the integration was assessed only in adult listeners and on a very small sample size.

Khan, Salian, Rajashekhar and Dhamani (2008) studied normal hearing and cochlear implant children to see the impact of temporal envelop and fine structure on audiovisual integration. In addition, audiovisual integration was also studied by Wasim, Chavan, Deema and Dhamani (2010) in dyslexic, poor academic performers and normal reading children. Apart from these, there are no other studies, to our knowledge, which have actually evaluated the developmental changes of audiovisual integration in the Indian population.

With respect to the stimuli used for evaluating audiovisual integration, most of the times, stop consonants were used. Further, it was found that when the bilabial was paired with velar consonants, McGurk effect was found to be more intense (McGurk & MacDonald, 1976, Massaro, 1984; Massaro et al., 1986; Grant & Seitz 1998; Hockley & Polka, 1994; Sekiyama & Burnham, 2004; Dupont et al., 2005), which indicates that place of articulation plays a role in the integration process.

The voicing cue, one of the important features in speech perception, could also be perceived through visual modality via observation of lip movement. However, only few number of studies were found in this aspect. Schwartz, Berthommier and Savariaux (2004) compared identification scores for voicing in audio-only and audiovisual conditions. They used pre-voicing continuum from /ba/ to /pa/. They found improvement in perception of voicing when visual cue was added with audition. In the present study, we were interested to see, if these results could be obtained for participants across age groups using a non-continuum stimulus.

Hence, the aim of the present study was to explore the effect of age on audiovisual integration across six age groups (from age range 6 to 20 years), for male and female listeners, using the following conditions: Audio only; Visual only; Audiovisual congruent; and Audiovisual incongruent condition to test the audiovisual integration through McGurk phenomenon and audiovisual integration of voicing cue.

Method

The method consisted of two phases. In the first phase, selection of participants was done and in the second phase, the actual experiment to evaluate audiovisual Integration was carried out. Prior to the first phase, a written consent was taken from the caregivers or the participants, after briefing about the study, its objectives and duration of testing.

Phase I

Selection of Participants: In the first phase, selection of participants was done using a checklist and various tests. They are Screening Checklist for Auditory Processing (SCAP), Tumbling-'E' chart test for checking the visual acuity, routine audiological evaluation, Speech in Noise Test (SPIN) and Gap Detection Test (GDT).

Initially, a total of 200 participants were selected only based on the age. The participants' age ranged between 6 to 20 years. There were 170 children and 30 adults. All of them had normal speech and language skills and had no history of complaint of any neurological deficits.

Further selection was based on a series of screening and diagnostic tests to rule out any auditory processing difficulties and hearing difficulties. Depending on the age group, the tests administered for selection varied. Out of the 200 participants selected, only 150 met the inclusion criteria based on the test results. Thus, these 150 participants were enrolled for the second phase.

Screening checklist for Central Auditory Processing (SCAP): This check list was developed by Yathiraj and Mascarenhas (2002, 2004). This checklist was administered on 170 participants from different schools. The participants with scores less than 6 were considered to have passed the checklist, as scores less than 6 indicates no processing problems. Out of 170 subjects, 140 participants scored less than 6 and were taken up for visual acuity test. The rest (30 in number) of the participants who failed were eliminated from the study at this stage itself.

Vision acuity test: These 140 children, who passed SCAP, and 30 adult participants, were tested for visual acuity by using tumbling 'E' chart (Taylor, 1976). For children, this test was carried out in a quiet and distraction free room in the school, and for adults, it was carried out in a quiet and well illuminated audiometric room. All 170 participants (140 children and 30 adults) passed the visual acuity test. Hence, all of them underwent a routine audiological evaluation.

Routine audiological evaluation: Routine audiological evaluation was carried out in an acoustically treated room. Air conduction and bone conduction thresholds were established using modified Hughsen and Westlake (Carhart & Jerger, 1959) procedure. Speech audiometry was also carried out on all the participants. A calibrated two-channel Madsen (Orbiter-922) audiometer with TDH-39 headphones was used to establish air conduction pure tone thresholds and speech audiometry. B-71 bone vibrator was used to establish bone conduction thresholds. Hearing was considered normal if puretone thresholds were within 15 dB HL bilaterally at all octave frequencies from 250 Hz to 8 kHz. Tympanometry and acoustic reflex thresholds were established using a calibrated Grason Stadler-Tympstar middle ear analyzer. Presence of 'A' or 'As' type of tympanogram with reflexes present in both the ears below 100 dB HL at 500 Hz, 1 kHz and 2 kHz was considered as normal.

Out of 170, 163 participants showed normal findings in routine evaluation. Seven participants were found to have abnormal type of tympanogram, ('C', 'Cs' or 'B') with reflexes absent and had poor hearing thresholds. These participants were excluded from the study.

For the participants who passed routine hearing test, SPIN and GDT were administered except for 15 children. These 15 children were below seven years of age. SPIN could not be administered on these children, as they had difficulty understanding instructions. GDT also could not be administered on them, as it has been standardized only for children above seven years of age. Thus, for these 15 children, between 6 to 7 years of age, the results of SCAP and routine audiological evaluation only were considered for inclusion. The remaining 148 (163 minus 15) subjects were considered for SPIN and GDT testing.

Speech in noise test: Speech in noise test (SPIN) was carried out for 148 participants. The phonemically balanced (PB) Kannada word list developed by Yathiraj and Vijayalakshmi (2005) was used for this. Speech identification scores in quiet and in the presence of ipsilateral noise were found out. The level of test stimulus and noise were 40 dB SL (ref. SRT).

Scoring was done by calculating percent correct responses. The difference between percent correct response in quiet and that in noise were then calculated. If speech identification scores in the presence of ipsilateral noise were poorer by 40% or more when compared to that of quite condition, it was considered as abnormal. All the 148 participants obtained normal SPIN scores.

Gap detection test: Gap detection test (GDT), developed by Shivprakash (2003) was administered on these 148 participants, to further confirm normal auditory processing abilities. This test consists of 56 trials, including 8 catch trials of broad band noise which contained gaps. The participants were required to indicate as to which set of noise bursts in a triad contained a gap. The minimum gap that the participant could detect was considered as gap detection threshold. This was compared with the normative given.

Out of the 148 participants, 13 of them obtained abnormal gap detection threshold. Hence, the remaining 135 participants, along with the 15 children below 7 years of age were included in the second phase of the study. Hence, there were a total of 150 participants for the next phase. These 150 participants were then divided into 6 groups based on the age (in years), with 25 participants in each group. The details of different groups and age are given in the Table 1.

Phase II

Measurement of audiovisual integration: In the second phase, the actual experiment of measurement of audio-visual integration was carried out using the following four conditions: Audio only (AO), Visual only (VO), Audiovisual congruent (AV+) and finally, Audiovisual incongruent (AV-).

AO and VO were unimodal conditions, and AV+ and AV- were bimodal conditions. In AV+ condition, the sounds presented through auditory and visual mode were the same, wherein, the AV- condition had conflicting auditory and visual stimuli. *Stimulus recording:* The stimuli used were six monosyllables, /pa/, /ba/, /ta/, /da/, /ka/ and /ga/. All the sounds were stop consonants with the vowel /a/. These CV monosyllables were uttered by a 26 years old female native Kannada speaker. The recording was done in a sound proof audiometric room. Audio and video recording of these syllables were done. The speaker was seated at a distance of six feet from the camera with the head and neck held erect. The auditory and visual stimulus was recorded simultaneously using a National m-7 movie camera with an inbuilt microphone. A TV zoom lens with the power of 1:12 was used, and 1 KW halogen light was used to illuminate the room.

The participants were asked to practice uttering the syllables before recording. During the recording, the participants were instructed to utter the monosyllables, four times each with a pause in between each of them. These were video-recorded on the video track of the cassette national VHS (spe-180). The cassette version of the video was then converted in to avi movie file, using the software 'Any video converter' for easy editing of the stimulus.

Stimulus editing: These avi files were edited further using Virtual dub (version 1.0). Based on the clarity of the video and the naturalness of the articulatory movements of the speaker, three best recordings out of the four recordings were selected for further editing. These stimuli were separated into audio and video files and were digitized using Adobe audition (Version 3). Video digitizing was done at 29.97 frames/s in 640 $\times$ 480 pixels, and audio digitizing was done at 44 kHz in 32 bits. All the syllables were normalized in order to avoid

Table 1: Details of different groups, Mean age and Standard deviation

Group number	Age range(in years)	Mean age (in years)	Standard deviation (SD)
I	6 -7.11	6.92	0.38
II	8 -9.11	8.88	0.52
II	10 -11.11	10.70	1.95
IV	12 -13.11	13.01	0.66
V	14 -15.11	14.93	0.51
VI	16 -20	18.60	1.11

the effect of intensity difference between different syllables.

Using these edited audio and video files, stimulus for AO, VO, and AV (+/-) conditions was generated. In AV condition, the audio and video stimuli were synchronized. As mentioned earlier for AV+ condition, the monosyllables in audio as well as video were same, where as in the AV- condition, monosyllables in the two modalities (audio and video) were different. For studying the presence of McGurk effect, auditory /pa/ was dubbed onto the articulatory movements for /ka/, and auditory /ba/ was dubbed onto the articulatory movements for /ga/.

For assessment of audiovisual integration of voicing cue, the voiced and voiceless sounds were paired (for e.g., audio voiceless /pa/ was paired with video of the /ba/ and vice versa). This was done for all the six stop consonants. The details about the stimuli in each of these conditions are given in the Table 2.

Goodness test: The prepared stimuli were presented to ten adults in the age range of 16 to 30 years. The participants were asked to rate in a three point rating scale which included: 1) Clear and good quality; 2) Distorted but can be identified; and 3) Distorted and cannot be identified.

Out of three utterances, the utterance which was rated as clear was taken up for making final stimulus. In addition, the stimulus prepared for assessing the McGurk effect, that is, Aud /pa/ - Vis /ka/ and Aud /ba/ - Vis /ga/ were also presented to those participants. The pilot study revealed that the Mc Gurk effect was weaker for the stimulus pair Aud /ba/ - Vis /ga/ than the other pair. Hence, only Aud /pa - Vis /ka/ was included in the study.

Seating arrangement: Each participant was tested individually in a double room situation. Participants were seated in a sound treated room comfortably in a chair. A laptop was placed 1 meter away from the participant on the table, in an appropriate height, so as to facilitate easy and normal access of the laptop screen (video) for the participants. A Martin Audio loud speaker was placed at 0 degree azimuth at the ear level of the participants.

Stimulus presentation: Stimuli were presented at 40 dB SL (ref. SRT). In all the conditions, the task was to repeat the syllables presented in auditory and/or visual modality. The responses were open set responses. In each condition, the task was preceded by a practice phase and a short break was given, if necessary. This was most often necessary for the children in the Group I (age range 6 to 7.11 years). The children above this age group were not given any repetitions of the instruction or the task.

Table 2: The stimuli used in different conditions

$AudioonlyAO$	$VisualonlyVO$	$Audiovisual(congruent)AV+$	$Audiovisual(incongruent)AV-ForMcGurkeffect$	$Audiovisual(incongruent)AV-Forvoicing$
$Auditory$	$Visual$	$Auditory-Visual$	$Auditory-Visual$	$AuditoryVisual$
$Aud/ba/$	$Vis/ba/$	$Aud/ba/+vis/ba/$	$Aud/pa/+vis/ka/$	$Aud/pa/+vis/ba/$
$Aud/da/$	$Vis/da/$	$Aud/da/+vis/da/$		$Aud/ta/+vis/da/$
$Aud/ga/$	$Vis/ga/$	$Aud/ga/+vis/ga/$		$Aud/ka/+vis/ga/$
$Aud/pa/$	$Vis/pa/$	$Aud/pa/+vis/pa/$		$Aud/ba/+vis/pa/$
$Aud/ta/$	$Vis/ta/$	$Aud/ta/+vis/ta/$		$Aud/da/+vis/ta/$
$Aud/ka/$	$Vis/ka/$	$Aud/ka/+vis/ka/$		$Aud/ga/+vis/ka/$

The stimuli in the unimodal conditions were presented first, i. e, AO and VO conditions. After three hrs of break, testing was done in the bimodal conditions (AV+, AV-). This gap of three hrs was given to avoid the memorization of the stimulus by the subjects. The responses were noted by an experimenter (who remained behind the participants). After each condition participants were instructed depending on the next task. The experiment (testing time) lasted for about 15-20 minutes all together.

Response elicitation and scoring: The participants were asked to give an oral response of what he/she had perceived. As previously explained, each stimulus was presented three times. If two responses out of the three presentation were same, that was taken as final response, regardless of whether it was repeated correctly or not. The number of CV monosyllables responded correctly for each individual participant, in different conditions, was calculated.

Results and Discussion

The aim of the present study was to evaluate the developmental trend of audiovisual integration across six age groups. The results are discussed under the following four headings: Auditory only (AO); Visual only (VO); Audiovisual congruent (AV+); and Audiovisual incongruent (AV-) condition.

Audio-only Condition

In this condition, signal was presented only through auditory mode and the participants were made to listen and speak what they heard. This auditory-only perception was assessed for all six sounds. The number of CV syllables perceived correctly (that is, out of six syllables presented, how many were repeated corrected) by each individual was calculated for each group.

From the Table 3, it can be observed that the auditory-only performance is not very different across different age groups and gender, though the younger groups have performed a little poorer than the older groups.

In order to see, if there is a statistically significant difference across groups and the gender, two way MANOVA was done. The analysis revealed no significant difference among different age groups [$F (5,138) = 2.349$, $p>0.05$] and between male and female participants [$F (1,138) = 0.462$, $p>0.05$] in the AO condition.

The above results show that participants of all age groups, that is, from 6 to 20 years of age, were able to perform equally when only auditory information was delivered. These results are consistent with the findings that the discrimination of the stop consonants have been observed in infants as young as 20 to 24 weeks of age (Moffitt, 1971). In the present study, all the participants were well above this age.

Visual only Condition

These were the responses of the participants for the unimodal visual only task. Here the participants were given only the visual stimuli (6 in number) and were instructed to give an oral response of what they saw. For each individual, the number of correctly identified CV syllables was counted. The mean and standard deviation of the number of syllables correctly identified for all the groups and for both the gender are given in Table 4.

It can be observed from the Table 4 that the Group I (6 to 7.11 year olds) performed higher than all the other groups. Group VI (16 to 20 year olds) got the second highest scores (though the SD of this group is smaller than the Group I) followed by Group II. Groups III (10 to 11.11 year olds), IV (12 to 13.11 year olds) and V (14 to 15.11 year olds) performed poorer than the other three groups.

Two-way MANOVA was used to determine if there was a statistically significant difference across groups. Results of two way MANOVA revealed a significant difference across groups [$F (5,138) = 8.251$, $p < 0.05$].

In summary, for VO condition, individuals between 6 to 7;11 years of age and 16 - 20 years of age had better performance. Individuals with 8 to 15;11 years of age performed poorer than other groups.

These results are not in agreement with other studies (McGurk & MacDonald, 1976; Desjardins et al., 1997;

Table 3: Mean and SD of number of correctly identified syllables across different age groups and gender for AO condition

Group	Age range (in yrs)	N	Mean for male participants	SD	Mean for female participants	SD	Total Mean	SD
I	6 to 7; 11	25	5.33	1.56	6.00	0.00	5.68	1.10
II	8 to 9;11	25	5.80	0.577	5.54	0.87	5.68	0.74
III	10 to 11;11	25	5.50	0.90	5.23	1.01	5.36	0.95
IV	12 to 13;11	25	6.00	0.00	6.00	0.00	6.00	0.00
V	14 to 15;11	25	5.58	1.44	6.00	0.00	5.80	1.00
VI	16 to 20	25	6.00	0.00	6.00	0.00	6.00	0.00

$N = no.of participants.$

Table 4: Mean and SD of number of correctly identified syllables across different age groups and gender for VO condition.

Group	Age range(in years)	N	Mean for male	SD	Mean for female	SD	Total mean	SD
I	6 to 7;11	25	4.92	0.96	3.84	0.61	4.36	1.65
II	8 to 9;11	25	3.08	1.44	3.08	1.48	3.44	1.47
III	10 to 11;11	25	3.08	1.78	2.92	1.25	3.00	1.50
IV	12 to 13;11	25	2.07	1.04	2.50	0.90	2.28	0.97
V	14 to 15;11	25	2.58	1.93	2.31	1.49	2.44	1.68
VI	16 to 20	25	3.33	0.77	3.23	1.57	3.84	1.06

Note: Here 'N' represents the no. of participants, 'SD' represents, standard deviation. Maximum no. of syllables- 6.

Table 5: Mean of number of correctly identified syllables across different age groups and gender in AV+ condition

Group	Age range (in years)	N	Mean for male	Mean for female	Total mean
I	6 - 7.11	25	6	6	6
II	8 - 9.11	25	6	6	6
III	10 - 11.11	25	6	6	6
IV	12 - 13.11	25	6	6	6
V	14 - 15.11	25	6	6	6
VI	16 - 20	25	6	6	6

Note: Here ?N? represents the no. of participants,. Maximum no. of syllables- 6.

Dupont, Aubin, & Menard, 2005; Hockley & Polka, 1994; Massaro, 1984; Massaro et al., 1986; Sekiyama & Burnham, 2004; Wightman et al., 2006), in which it has been found that the visual influence is lesser in the younger age and hence, younger children are poorer speech readers when compared to adults.

In the present study, the better performance of the youngest age group, i.e., 6 to 7;11 years reflects good visual processing. Further, the children in this group were repeatedly instructed to pay attention which might have added to the better performance. In addition, the equal responses seen for children of 8 to 15;11 years may be because of the low interest and low attention during the testing. Probably during the testing, for this age group attention to each item was not monitored carefully, whereas it was done for the younger group. Adults do not require such monitoring and hence the performance was good.

Comparison between the performance of the female and male listeners was made and the interaction between the group and gender interaction were analyzed. From the Table 4, it can be observed that between male and female participants there is not much of difference seen. Results of two way MANOVA revealed no significant differences for gender [$F(1,138) = 0.175$, $p > 0.05$].

Audiovisual Congruent Condition

In this condition response for the congruent stimuli pairs were recorded. The stimuli in auditory and the visual modality were same in terms of place of articulation and voicing in AV+ condition. The participants responded to the stimulus as a whole, regardless of any specific modality.

Analysis of the sounds perceived correctly by each individual for all the groups was done. The results of this are given in the Table 5' should be added at this point. It is shown that all the groups performed equally and there was 100% responses. These findings can be attributed to the better access of auditory and visual information when presented simultaneously. It is also observed that the mean value of this condition is higher than that of the AO condition. This is a clear cut indication of increase in performance because of addition of visual information.

This is in agreement with the results of previous studies, which revealed that the addition of the visual cues to an auditory signal facilitates/ enhances overall perception of speech sound significantly (Sumby & Pollack 1954; Erber, 1969). Further, these were no gender differences seen in this condition.

Table 6: Fused and nonfused responses for the stimulus 'Aud /pa/ and Vis /ka/' (In each group, n=25)

Group	Total no. of participants perceived /ta/ Fused response	Total no. of participants perceived /pa/ Nonfused response	Total no. of participants perceived /ka/ Nonfused response
I	17	4	4
II	7	14	4
III	21	2	2
IV	20	3	2
V	20	3	2
VI	21	0	2

Note: Here N= Total number of participants 'Aud' represents auditory mode and 'Vis' represents visual mode.

Audiovisual Incongruent Condition

For studying the presence of McGurk effect, auditory /pa/ dubbed onto the articulatory movements for /ka/ was presented. McGurk effect was said to be present, if there was a fused response, that is, /ta/ in response to auditory /pa/ and visual /ka/ stimuli. McGurk effect was said to be absent, if the participants perceived either auditory or visual stimulus, i.e., /p/ or /k/ respectively, instead of /ta/. If the participants' perceived /pa/, his/her perception was said to be auditory biased. If he/she perceived /ka/, his/her perception was said to be visual biased.

The total number of participants who perceived /ta/ and the total number of participants who perceived either /pa/ (auditory biased response) or /ka/ (visual biased response) were calculated. In the Table 6, the details of these responses across different groups are given.

From the Table 6, it can be observed that the performance of group II is considerably lower than the other groups. It can also be observed that the performance of all the other groups did not differ from each other, except for the Group II, which has relatively poorer performance when compared to Groups III, IV, V and VI. Test of equality of proportion using Smith's statistical package (SSP) software was performed in order to find out if there is a significant difference across groups. Only the pairs of groups between which there was significant difference were given in the Table 7.

Table 8: No. of male and female participants who gave fused responses across groups

Table 7 shows that there is a statistically significant difference in the performance of group II when compared to all the other groups. That is, 8 to 9;11 year olds, performed significantly poorer than that of children in the age range of 6 to7;11 years and that of 9 years and older children. Further, there was no significant difference between all the other groups ($p>0.05$), though the performance of the youngest group was a little poorer when compared to Group III, IV, V and VI. That is, the performance of 6 to 7.11 year olds and of children who were $\geq$ 9 years of age did not differ significantly.

This is contradicting to the results obtained in the studies (McGurk & MacDonald 1976; Rosenblum et al., 1997;, Desjardins et al., 1997; Dupont et al., 2005; Hockley & Polka, 1994; Massaro, 1984; Massaro et al., 1986; Sekiyama & Burnham, 2004; Wightman et al 2006) where, they have found poorer integration for children below 11years of age when compared to adults.

This may be because of the good VO scores by 6 to 7;11-year-olds in the visual only task. The innate differences of articulatory experience of the Indian children when compared to English speaking children may also be a reason. The support for this may be taken from the finding that the phonological structure of the language influences the audiovisual integration (Sekiyama & Burnham, 2008). It may also be that, the children taken in the present study had good phonological skills. However, this has to be tested for giving conclusive statement.

Analysis of the number of subjects who perceived /pa/ and who perceived /ka/, instead of a fused response, was also done. Table 6 provides the details of this. It is shown in the table that, only seven of 25 of the participants in Group II perceived /ta/, i.e., fused response. Most of the other participants in this group (14 in num-

Table 7: Pairwise comparison of fused responses

Stimulus Aud-Vis pa-ka Pair wise comparison					
Group	I-II	II-III	II-IV	II-V	II-VI
Z value	2.83	3.98	3.68	3.68	3.98.
P value	0.004**	0.000**	0.000**	0.000**	0.000**

*Note: Here ** depicts the significant difference between groups at the level of $p<0.01$*

Table 8: No. of male and female participants who gave fused responses across groups

Fused response for Pa-ka stimulus			
Group	Total no. of participants given fused responses	Total no. of female who gave fused responses	Total no. of male who gave fused responses
I	17	0	8
II	7	5	6
III	21	8	7
IV	20	9	6
V	20	11	5
VI	21	13	6

Note: Here the total no. of participants will not be the sum of male and female participants.

Table 9: Within group gender differences in the perception of fused response.

Stimulus pa - ka	Within group gender comparison for fused responses		
Group	I	V	VI
Z value	3.23	2.76	2.92
P value	0.001**	0.005**	0.003**

Note: * *significant difference at the level of p¡ 0.01*

ber), perceived /pa/ and four of them perceived /ka/. It is also shown in the table that, in all the other groups, almost equal number of participants perceived /pa/ and /ka/.

Hence, the test of equality of proportions was done only for Group II to check if there was a statistically significant difference between the auditory and visual response. It was found that there was a significant difference ($p<0.01$). This implies that, in 8 to 9;11 year olds, the auditory influence was more and they had significantly poorer audiovisual integration compared to all others.

This is in agreement with the results of the study done by Erderner and Burnham (2005). They found that there was a decline in the visual speech influence at around 7 years of age. Further, at the age of 8 years their integration jumped back to that of 6-years -old and their reading skill also was found to be increased substantially.

The reason attributed was that the reading skill starts as an automatic skill at 7 years of age. Hence, visual speech perception is not needed as a back support. Because of this they performed poorly.

However, in the present study, the performance drop was between 8 to 9.11 years old children. Hence, we further analyzed the age of seven participants who gave fused response in this group. It was noticed that these seven children were of 9 years or older. All the others (who did not give fused responses) in the group were younger (around 8.2 years) than those seven children. Hence, children at around 8 years of age showed poorer integration in contrast to (at 7 years of age) Erdener and Burnham's study.

The reason for this difference may be that, in the present study, all the children had Kannada language as mother tongue (communicate in Kannada language at home) and English as the medium of instruction at school.

Given this and the complexity of phoneme-grapheme correspondence of English language, the children of 8 years of age, in the present study, might have just developed reading as an automatic process. However, this is just a hypothesis, which needs to be tested.

Another objective of the study was to look for gender differences. The responses of male and female participants are given in Table 8. It was observed that more number of female participants have given fused responses when compared to male participants, except for the Group I. In the Group I, none of the female subjects gave fused responses and eight of male participants gave fused response. The test of equality of proportion was done in order to compare the performance between male and female participants across different age groups. It was found that, only for group I, V and VI, there was a significant difference between male and female participants, as given in Table 9. Females performed better when compared to male participants, except for group I where male did significantly better than

Table 10: Mean and SD of the numbers of syllables perceived across age groups in the audiovisual integration of voicing cue task

Group	Age range (in years)	N	MeanFor audio	SD	Meanfor video	SD
I	6 to 7;11	25	6.00	0.00	0.00	0.00
II	8 to 9;11	25	6.00	0.00	0.00	0.00
III	10 to 11;11	25	6.00	0.00	0.00	0.00
IV	12 to 13;11	25	6.00	0.00	0.00	0.00
V	14 to 15;11	25	6.00	0.00	0.00	0.00
VI	16 to 20	25	6.00	0.00	0.00	0.00

Note: 'N' represents no. of participants in each group; 'SD' represents standard deviation.

females. In the other groups, there were no significant gender differences found.

In most of the studies, female subjects were found to show a stronger McGurk effect than males (Iyer, 2005). This was attributed to female listeners' ability to have superior lip reading abilities and was better at integrating audiovisual speech information than the male listeners (Johnson et. al. 1988; Iyer, 2005).

This is in agreement with the findings of the present study for children of 14 years of age and above. However, in three of six groups, there were no gender differences. This conflicting result may be because of the methodological differences with Iyer's study, in which integration was assessed in noisy condition.

Measurement of Audiovisual Integration for Voicing

In this condition, the voiced and voiceless sounds were paired, for example, audio voiceless /pa/ was paired with the video of /ba/ and vice versa. These were a total of six such pairs. The performance was checked to find out if there is an influence of visual stimulus in the perception of voicing when an incongruent auditory stimulus is given.

It can be observed from the Table 10 that when an auditory stimulus was voiced, participants perceived a voiced stimulus and when it was voiceless, participant perceived a voiceless sound. Because there was 100% response given by all the groups and the responses were same, no statistical analysis was done. All the groups repeated the auditory stimuli and there was no influence of vision. The findings of the present study are not in consonance with the previous studies (Schwartz et al., 2004).

Conclusions

To conclude, the developmental changes of audiovisual integration of the Indian children is different from the developmental changes reported in western studies. Further, it can be stated that the assessment of phonology, reading and writing skills of the children whilst looking into the audiovisual integration will give better insight about the process by itself. This study again confirms that the visual information facilities speech perception in the congruent condition. It can also be concluded that it is important to study the female and male participants separately and, also, to have a task which is more demanding such as testing in the presence of noise.

The developmental changes of audiovisual integration of the Indian children are different from the developmental changes reported in western studies. This demands more research to include cross linguistic studies in the Indian population. Secondly this study is one of the first of its kind in Indian population and the results of the study help in providing baseline for future research. Finally, it also adds onto the literature of audiovisual integration.

Acknowledgement

I would like to thank Prof. S.R. Savithri director of All India institute of Speech and hearing, Mysore and HOD of Audiology Dr. Animesh Barman who allowed carrying out the research and also to all the participants and their parents for their cooperation.

References

Carhart. R., & Jerger. J. (1959). Preferred Method of clinical Determination of Pure Tone Threshold. *Journal of speech hearing Disorder,24*, 330-345.

Conrad, R. (1977). Lipreading by deaf and hearing children. *British Journal of Educational Psychology, 47*, 60-65.

Dodd, B. (1977). The role of vision in the perception of speech. *Perception, 6*, 31-40.

Dodd, B. (1980). Interaction of auditory and visual information in speech perception. *British Journal of Psychology, 71*, 541-549.

Desjardins, R., Rogers, J., & Werker, J. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology, 66*, 85-110.

Dupont, S., Aubin, J., & Menard, L. (2005). A study of the McGurk effect in 4- and 5-year-old French Canadian children. *ZAS Papers in Linguistics, 40*, 1-17.

Erber, N. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research, 12*, 423-425.

Erdener, V. D., & Burnham, D. K. (2005). Development of auditory-visual speech perception in English-speaking children: The role of language-specific factors. In E. Vatikiotis-Bateson, D. Burnham & S. Fels (Eds.), *Proceedings of Auditory-Visual Speech Processing International Conference* (pp.57-62). Vancouver Island, BC, Canada.

Grant, K. W., & Seitz, P. F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *Journal of the Acoustical Society of America, 104*(4), 2438-2450.

Green, K. P., & Kuhl, P. K. (1991). Integral processing of visual place and auditory voicing information during phonetic perception. *Journal of Experimental Psychology, 17*, 278-288.

Hockley, N., & Polka, L. (1994). A developmental study of audiovisual speech perception using the McGurk paradigm. *Journal of the Acoustical Society of America, 96*, 3309.

Iyer K. (2005). *Audiovisual Speech Perception Abilities of Children with Normal Hearing in Quiet and in Noise & Comparison to Participants with Cochlear Implants.* Unpublished Master of Speech language therapy research project. University of Auckland.

Jerger, J., Jerger, S., Oliver, T., & Pirozzolo, F. (1989). Speech understanding in the elderly. *Ear and Hearing, 10*, 79-89.

Khan, M. K. F. Salian. T. S., Rajashekhar. B. & Dhamani. I., (2008). *Effect of Temporal Envelope and Fine Structure cues in Auditory-Visual Integration in children using Cochlear Implants*. Unpublished oral paper presented in CIGICON, Bangalore.

Massaro, D. W. (1984). Children's perception of visual and auditory speech. *Child Development, 55*, 1777-1788.

Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, 41*, 93-113.

McGurk, J., & MacDonald. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

Moffitt, A. (1971). Consonant cue perception by 20 to 24 weeks old infants. *Child development, 42*, 717-731.

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception and Psychophysics, 59*, 347-357.

Sarika, M. (1995), *Cross- model perception of speech reading and auditory mode*. Unpublished Master's Dissertation, University of Mysore, Mysore.

Schwartz, J. L., Berthommier, F., & Savariaux, C., (2004). Seeing to hear better: Evidence for early audiovisual interactions in speech identification. *Cognition, 93*, 69-78.

Sekiyama, K., & Burnham, D. (2004). Issues in the development of auditory-visual speech perception: Adults, infants, and children. *Interspeech, 21*, 1137-1140.

Shivprakash, S. (2003). *Gap Detection Test - development of Norm*. Unpublished Master's dissertation. University of Mysore, Mysore.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society America, 26*, 212-215.

Taylor, H. R. (1978). " Applying New Design Principles to the Construction of an Illiterate E Chart." *American Journal of Optometry and Physiological Optics, 55*, 348.

Wasim, A., Chavan. A. R., Deema, K. P. A., & Dhamani, I. (2010). *Auditory visual integration and dyslexia and poor academic performers*. An unpublished Poster paper presentation in ISHACON.

Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *Journal of the Acoustical Society of America, 119*, 3940-3949.

Yathiraj, A., & Mascarenhas. K. (2002). Audiological profile of the children with suspected auditory processing disorder. Developed as a part of project titled *'Effect of auditory stimulation of central auditory processes in children with CAPD*' at the department of audiology. All India Institute of Speech and Hearing, Mysore.

Yathiraj, A., & Mascarenhas, K. (2004). Auditory profile of children with suspected auditory processing disorder. *Journal of Indian Speech and Hearing 18,* 6-14.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). Phonemically Balanced word list in Kannada, Developed in Department of Audiology, AIISH, Mysore.

Efficacy of a Hearing Checklist and Screening Test in Identifying Hearing Problems in Primary School Children

[1]Ratul Dey & [2]Asha Yathiraj

Abstract

The study evaluated the sensitivity, specificity as well as the positive and negative predictive values of a hearing screening checklist and a screening test developed using the Ling's 6-sound test. Hundred and fifty-four primary school children, aged 5 to 7 years, were screened and later tested on standard audiological tests (diagnostic pure-tone audiometry, tympanometry, acoustic reflex test & DPOAE). The screening checklist 'Signs and symptoms of hearing loss' was found not to be effective in detecting the presence of hearing impairment. The recorded 'Ling 6-sound screening test' was found to be a better predictor of the presence of hearing loss compared to the screening checklist as it had a higher sensitivity and specificity. The screening test also had relatively low over referral and under referral rates when compared to the checklist. The combination of the checklist and the screening test did not give rise to significant improvement in the effectiveness of the screening program. Hence, the use of the Ling's 6 sound screening test in isolation is recommended.

Keywords: *School screening test, symptoms of hearing loss, sensitivity, specificity.*

Introduction

Hearing is known to play a vital role in the acquisition of speech and language as well as in the achievement of other developmental milestones in young children. Undetected hearing impairment, especially in children, has been found to cause lifelong disturbance in social, emotional, behavioural and cognitive spheres or combination of any of them (Yoshinaga-Itano, Sedey, Coulter & Mehl, 1998; Yoshinaga-Itano, 2003). Undetected hearing loss has also been found to negatively impact the educational achievement of children (Nix, 1977; Northcott, 1972; Maxon & Brackett, 1981).

The prevalence of hearing loss in school-age population was found to be 11.3% by Bess, Dodd-Murphy and Parker (1998). The study also reported that 9.62% to 12% of school-going children were at-risk for hearing loss. In another study by Olusanya (2001), the prevalence of hearing loss was found to be 13.9 % out of which 3.3 % was sensorineural hearing loss. In Mysore city of India, Nikam and Dharamraj (1971) found the incidence of hearing loss to be 3.9% in children aged 2 to 14 years. Further, the report of the National Sample Survey Organisation (NSSO), Government of India (1991), showed that the prevalence of hearing impairment in the age group of 0 to 14 years was 2.7% in rural India and 3% in urban India. An increase in the prevalence of hearing impairment was reported by NSSO (2002), where it was found to be 4.23% and 4.06% for rural and urban area respectively. The Human Development report of 1999, estimated a 0.3 million hearing impaired population between the ages of 0 to 4 years and 1.5 million in the age range of 5 to 12 years in India.

A major concern regarding hearing loss in children is that, in many cases it is identified only at a later stage based on symptoms. By this time the long term consequences of hearing loss have been noted to have already occurred (Nozza, Sabo & Mandel, 1997). Therefore it is important to detect the presence of it as early as possible for better intervention. In order to identify hearing impairment early, hearing screening has been recommended. The practice of hearing screening programmes in schools has been reported to have started in the 1920s but became a routine part only by 1960s (Northern & Downs, 2002). Harford, Bess, Bluestone and Klein (1978) described hearing screening as rapid and simple measures that identify those individuals who have a high possibility of a disorder which may otherwise go undetected. According to Alpiner (1976), hearing screening programmes include early identification of children through audiological screening and medical evaluation. Based on this, further recommendations for rehabilitation and periodic follow-up evaluation were made. Such a program was also noted to enable awareness about prevention of hearing loss, planning for appropriate rehabilitation and educational programmes for children with significant amount of hearing loss.

Due to the negative impact of hearing impairment, it should be mandatory to run hearing screening programmes routinely among school-age population with the best available and suitable tools and tests. Teachers can be the key persons to the whole process. Training them with a suitable screening test may enable early identification of such population.

The use of only checklists in identifying hearing problems in school children has been noted to have low sensitivity (Curry, 1950; Kumar & D'Mello, 2006). Hence, the tests such as pure-tone hearing screening have been used extensively (Glorig & House, 1957; Norton & Lux, 1960; FitzZaland & Zink, 1984; ASHA, 1997;

[1]Email:deysworld@gmail.com
[2]Professor of Audiology, Email:asha_ yathiraj@rediffmail.com

Niskar, et al.,1998; Sarafraz & Ahmadi, 2009). However, it has been observed, that pure-tones are novel stimuli, which children are not exposed to. Martin (1991) reported that this has resulted in children finding them difficult to respond, thus increasing the false positives. On account of this difficulty, the addition of objective measures such as immittance audiometry has been advocated for school screening (Paradise & Smith, 1979; Krueger & Ferguson, 2002; Lewis, Dugdale, Canty & Jerger, 1975; FitzZaland & Zink 1984). This addition was found to improve the sensitivity (Krueger & Ferguson, 2002; Lewis et al., 1975). However, this would increase the cost of school screening considerably. In view of these issues, in addition to using checklists it is necessary to use a comprehensive screening test that can be used easily without the need for expensive specialized equipment. Thus, in a country like India, there is a need for effective, quick, simple and cost effective school screening procedures, due to the large population.

Speech materials as screening tool have not been advocated due to its negative effects (Ritchie & Merklein, 1972). Mencher and McCulloch (1970) attributed this to the audibility of some high intensity phonemes in some words which could act as cue in a few frequencies in children with mild hearing loss. However, this can be overcome with the use of selected speech sounds covering the speech spectrum and recorded to control intensity variations. Therefore, the use of speech sounds as a school screening procedure should be considered, especially since they are familiar to children.

The importance of hearing screening in India is documented in the 'Persons with Disabilities (Equal opportunities, Protection of Right and Full Participation) Act' (1995). The Act stipulates the necessity for early identification and intervention. Further, the National Programme in Prevention and Control of Deafness strongly promotes early identify hearing impairment in children. In order to carry out such a task, it is essential to have a cost effective as well as time effective hearing screening procedure that can be easily used in any part of the country. It is also important to study the effectiveness of such a screening procedure. Thus, the present study aimed to check the sensitivity and specificity of a hearing screening checklist and a pre-recorded hearing screening test when administered alone and when combined.

Method

The study was conducted in three phases. Phase-I involved the administration of a checklist 'Signs and Symptoms of Hearing loss' by the class teachers; Phase-II involved screening children with a recorded version of the Ling's 6-sound hearing screening test; and Phase-III involved running standard hearing screening tests and a diagnostic test.

Participants

One hundred and fifty-four primary school children aged 5 to 7 years from 5 schools were evaluated. Children studying in Grade-I or Grade-II were selected randomly for the study. Among the participants, 94 were from Kannada medium schools and 60 were from English medium schools. All the children spoke Kannada or English fluently. None of the children had any articulation problem and all of them were able to clearly produce the Ling's speech sounds (/a/, /i/, /u/, /m/, /s/ and /sh/), as reported by the teacher. On the day of the screening none of the children had any illness.

Further, 10 school teachers who taught the participants, were required to provide information regarding the hearing abilities of the children. Only those teachers, who had taught the children for 6 months or more, provided this information.

Material:

Two different school screening material were used to determined their sensitivity and specificity. The hearing checklist 'Signs and Symptoms of Hearing loss' developed in the Department of Audiology, All India Institute of Speech and Hearing, Mysore, was used to obtain the responses from the class teacher regarding their opinion about the hearing abilities of the children. A recorded version of the 'Ling's-6 sound test', which was done as a part of the present study, served as the second screening procedure.

Recording of the Ling's-6 sound test: The material was recorded by a female volunteer having clear speech and normal fundamental frequency. The recording was done using a sampling rate of 44.1 kHz and 32-bit analogue-to-digital converter in an acoustically treated room. Praat (version: 5.1.31) software was used for the recording. A unidirectional microphone (Ahuja AUD 101XLR) was placed at a distance of 6 inches from the mouth of the speaker. The recorded material was later edited and scaled using Adobe Audition (Version: 1.5) software to ensure that the intensity of all sounds were at the same level.

A goodness test was run on 10 adults to confirm the quality of the recorded material. The recording was redone until it was confirmed by all 10 adults that each of the recorded speech sounds was intelligible and not distorted. Six different lists were made by randomizing the order of the 6 speech sounds. A 1 kHz calibration tone was added prior to each of the lists.

Further, a pilot study was done to determine the intensity level at which the stimuli should be played in an environment that simulated the noise levels of a typical quiet classroom. Prior to the pilot study, the output level of the recorded material from a laptop was determined using a sound level meter (Larson Davis 824) in a sound treated room. The volume control of the audio software and the computer were manipulated so that the output level through TDH-39 headphones was 25 dB HL and 30 dB HL. These levels were measured using an artificial ear (IEC318 Compliant Artificial Ear Coupler: Model AEC 101) with a 1 inch precision pressure response microphone (Model 2575). The control settings in the computer and the software for each of these intensity levels were noted. All the 10 adults and 10 children who participated in the pilot study could identify all six stimuli at 30 dB HL. However, only 40% of the adults and 20% of the children could identify the stimuli at 25 dB HL. They were able to identify the signals in an environment that simulated a quiet classroom, with activities going on outside. Based on these findings, it was decided that the screening should be done at 30 dB HL.

Test Environment

All the screening and diagnostic tests were done in quiet, well-lit rooms in schools. The rooms were large enough to accommodate the children and the examiners. The doors and windows were kept close to reduce the disturbance of noise emanating from outside the classroom. Sources of noise within the classroom, such as the CPU of a computer, fan, and other noise sources were switched off to ensure minimum noise interference. The ambient noise levels were such that a correction factor of not more than 15 dB had to be applied during testing. The thresholds of 3 normal hearing individuals, whose hearing thresholds had been earlier determined in a sound treated room that met the specification of ANSI S3. 1 (1991), were used for the biological calibration daily. Additionally, it was ensured that in each environment, Distortion Product Oto-Acoustic Emissions (DPOAEs) could be measured on all 3 normal hearing individuals, and that they could identify all the six stimuli of the recorded Ling's 6-sound test, when played at 30 dB HL.

Instrumentation

A calibrated diagnostic audiometer (Maico MA- 53) coupled with TDH-39 headphone and B-71 bone vibrator was used to estimate the pure-tone thresholds. Noise-excluding audio cups were used with the headphones. A calibrated middle ear analyzer (Interacoustics Titan) was used to carry out immittance tests. DPOAEs were evaluated using a calibrated OAE analyzer (GSI AUDIOscreener).

Procedure

As mentioned earlier, the study was conducted in three phases. The audiologist who evaluated the children during Phase-II, was blind to the findings of Phase-I. Similarly, a second audiologist who ran the standard screening tests and the diagnostic test was blind to the findings of Phase-I and Phase-II.

Phase-I: The school teachers were instructed to answer the questionnaire 'Signs and Symptoms of Hearing Loss' regarding the children in their class. Their doubts regarding the questionnaire were clarified in English or Kannada depending on the medium of instruction of the school. Children with a positive response on any of the questions of the checklist were categorised as 'refer' and those with no were categorised as 'pass'.

Phase-II: All the children who were administered the hearing checklist were also screened using the recorded Ling's 6-sound test in a quiet room. The test was played using Adobe Audition (version 1.5), with the volume controls of the software and the computer set such that the output from the TDH-39 earphones were 30 dB HL. Prior to placing the headphone on the children, they were seated comfortably and were instructed to repeat the speech sounds heard by them in each ear. Both ears of the children were tested independently using 2 of the 6 different lists of the Ling's 6-sound test. The screening was done first in the right ear for half the children and in the left for the other half to avoid any ear order effect. Children who could repeat all the sounds correctly were categorised as 'pass' and those who could not repeat one or more speech sounds were categorised as 'refer'.

Phase-III: All the children, both those who were marked 'pass' and 'refer', were later tested using standard screening tests. The standard screening tests consisted of immittance audiometry and DPOAE. Diagnostic pure-tone audiometry was also done to confirm whether the children had normal hearing thresholds.

Immittance evaluation was carried out for both ears of all the children. Tympanograms were obtained using standard 226 Hz probe tone and pressure sweep between +200 to -400 daPa. Ipsilateral acoustic reflex was recorded at 1000 Hz at 100 dB SPL. Those with 'A', 'As' and 'Ad' type tympanogram and reflex present were considered to be normal. Whereas 'B', 'C' and 'Cs' type tympanogram and absence of reflex were considered to be abnormal findings indicating pathological middle ear status. This criterion was considered based on the recommendation of FitzZaland and Zink (1984).

OAEs were also obtained for both the ears of all the children. DPOAEs were recorded using an f2: f1 ratio of 1.22: 1 and intensity level of 65 and 55 dB peak SPL (L1 and L2). The children were marked 'pass' and 'refer' based on the interpretation shown by the instrument. The instrument marked a child as 'pass' if 3 out of 5 frequencies were pass which met a minimum -5 dB SPL amplitude and a minimum 8 dB SNR, or a minimum noise floor amplitude of -17 dB SPL (NIH, 2000). Testing was repeated in cases of 'refer' to confirm the findings.

Pure-tone audiometry was administered after instructing the children in Kannada or English, depending on their fluency in that language. Children were seated facing away from the audiometer to prevent them from getting any visual cues. AC thresholds in the octave frequencies 250 Hz to 8 kHz and BC thresholds in octave frequencies from 250 Hz to 4 kHz were measured using the modified Hughson and Westlake procedure. From the pure tone thresholds, PTA 1 (average of thresholds at 500 Hz, 1 kHz & 2 kHz) and PTA 2 (average of thresholds at 1 kHz, 2 kHz & 4 kHz) were calculated. Children having both AC and BC thresholds within 15 dB HL were considered to have normal hearing sensitivity. Correction, based on the biological calibration values were applied to arrive at the actual thresholds of the children.

Test-retest reliability

Test-retest reliability was determined for results obtained in all 3 phases. To check the test-retest reliability of the responses to the checklist, 4 of the teachers (40%) were again asked to answer it for 16 (10.38%) of the children who were randomly selected. The screening and diagnostic tests were also re-administered on these 16 children. Re-administration of all the tests was done after a gap of 3 days.

Analyses

The obtained data was tabulated and based on the test findings, the participants were divided into two groups depending on whether they passed or were referred depending on the screening checklist and / or screening test. The referred children consisted of 3 groups: Children who were referred based on the checklist alone; Children who were referred based on the screening test alone; and children who were referred based on both the screening checklist and the screening test.

The sensitivity and specificity of the two trial screening procedures (the checklist 'Signs and Symptoms of Hearing Loss' and the screening test, 'Ling's 6-sound screening test') as well as the positive and negative predictive values were determined. This was done by comparing the findings of the two trial screening procedures with standard procedures (Immittance screening & DPOAE, and pure tone thresholds). Additionally, the test-retest reliability of the screening procedures was also determined.

The data were subjected to statistical analysis using SPSS software (Version: 20). Kappa test of agreement was used to compare the findings of each of the trial screening procedures (checklist & Ling's 6-sound screening test) with the results of the standard screening and diagnostic battery. Comparison was also done between the combined results of the two screening procedures and the diagnostic tests. Further, the relationship between the checklist and the screening test was also determined using Kappa test. Cronbach's alpha (α) reliability co-efficient was determined to check the test-retest reliability of all the tests.

Results and Discussion

The results of the 154 children (308 ears) are discussed under four broad headings: A. Findings of the hearing checklist; B. Findings of the Ling 6-sound screening test; C. Findings of the combination of hearing checklist & Ling 6-sound screening test; D. Test retest reliability measures. For the screening checklist and the screening test the following are provided: the general outcome, the sensitivity and specificity, the agreement with the standard tests, and the positive and negative predictive values.

Findings of the Hearing Checklist

General outcome of the hearing checklist: Of the 154 school children evaluated, 45 were referred based on the findings of the screening checklist. The symptom that occurred most frequently was, 'Is the child always distracted and is engrossed in other activities while the class is being conducted? ' The other symptoms that were present fairly frequently were, 'Does the child ask for frequent repetition of the message being spoken?', 'Does the child have an evident mispronunciation?' and 'Does the child have ear pain?' The symptoms that were observed more frequently did not directly relate to the presence of hearing loss but were indirect indicators of the problem. This highlights that teachers relied more on overt symptoms that may be pointers to detect a hearing impairment, rather than a direct indicator of hearing impairment.

Further, cluster analysis was done to subgroup these questions. It was found that the question could not be clustered into different subgroups. This indicated that each question in the checklist was unique. Based on this it is recommended that all the questions be utilized while administering the checklist as they tapped different aspects related to hearing.

Sensitivity and Specificity of the Screening Checklist: To find out the effectiveness of the screening checklist, its sensitivity and specificity was calculated. This was done using the following formulae:

$$\text{Sensitivity} = \frac{\text{No. of participants having hearing loss}}{\text{Total number having hearing loss}} \times 100\%$$

$$\text{Specificity} = \frac{\text{No. of participants not having hearing loss}}{\text{Total number not having hearing loss}} \times 100\%$$

The findings of the checklist, when compared to the different standard tests, showed a sensitivity ranging from 33.6% to 49.3% and specificity ranging from 72.8% to 76.4%. The sensitivity and specificity varied depending on the standard test with which it was compared. With tymanometry as the standard test, the sensitivity and specificity were 49.3% and 76.3%. Similar sensitivity and specificity were obtained with the standard tests, acoustic reflex (42.2% & 76.3% respectively) and DPOAE (42.4% & 76.4% respectively). Comparatively lesser sensitivity values were found when diagnostic PTA 1 (38.5%) and PTA 2 (36.7) were the standard tests though specificity did not vary much (72.8% & 74.9%).

Further, the Kappa measure of agreement was done to check the agreement between the checklist and different standard tests. It was found that there was a poor agreement of less than 0.228 between the findings of the screening checklist and that of the standard screening tests. However, this agreement was statistically significant ($p < 0.05$) for PTA 2, tympanometry, acoustic reflex and DPOAE.

Thus, from the findings of the sensitivity, specificity and Kappa measure of agreement, it can be inferred that the checklist was not effective in detecting the presence of hearing impairment. Similar findings have been reported in literature by Curry (1950), who found that only 7.4% out of all children who might have had a hearing loss were referred by school teachers.

Predictive values for the screening checklist: The positive predictive value (PPV) and negative predictive value (NPV) of the checklist were calculated to determine the over referral and under referral rates. PPV provided information regarding the ratio of the number of children detected correctly by the checklist as having a hearing problem to the total number of children who were positive on the checklist. Similarly, negative predictive value (NPV) calculated the ratio of the number of children who were detected correctly by the checklist to not have a hearing problem to the total number of those who were negative on the checklist. The values were calculated using the following formulae:

$$\text{PPV} = \frac{\text{No. identified correctly to have a problem}}{\text{Total number positive on the checklist}} \times 100\%$$

$$\text{NPV} = \frac{\text{No. identified correctly to not have a problem}}{\text{Total number negative on the checklist}} \times 100\%$$

The use of only the screening checklist resulted in unacceptably high over-referral (56.7% to 71.1%) and under-referral (68.3% to 84.4%) rates. The over-referral and under-referral rates varied depending of the standard test. Almost similar over-referral rates were found with PTA 2 (56%), acoustic reflex (56.7%), DPOAE (61.1%) and tympanometry (63.3%) as standard tests. A slightly higher over-referral rate was found when PTA 1 was considered as standard test. Almost similar under-referral rates were found with PTA 1 (78.4%), tympanometry (84.4%), acoustic reflex (75.7%) and DPOAE (78%) as standard tests. However, slightly lower under-referral rate (68.3%) was found with PTA 2. These results are in agreement with the findings of Olusanya (2001), who found that the sensitivity and specificity for a hearing screening questionnaire were 10 % and 94 %. In the same study a positive predictive value of 21.7% was found for the questionnaire. However, Lo, Tong, Wong and Hasselt (2006) in an effort to determining the accuracy of parental suspicion to detect hearing problem in children, found a sensitivity of 19.7% and specificity of 96.9 %. Further, they also got a positive predictive value of 82% and negative predictive value of 62.1%. The sensitivity reported by them is comparable with the findings of the present study. However, the positive (28.9% to 44%) and negative (15% to 31.7%) predictive values were comparatively lower in the present study. Their higher values were probably obtained because the responses were obtained from parents who were more sensitive to the presence of a hearing problem in their wards than school teachers. In the present study, the checklist was answered by the teachers who probably did not have the same level of sensitivity to the presence of a hearing problem as parents. This discrepancy can be attributed to the fact that parents spent more one-on-one time with the children and hence could detect the presence of a hearing problem. Another reason could have been that the parents had to make judgments about only a limited number of children whereas teachers had to make judgments about a much larger number of children.

The findings of the present study and that published in literature highlight that using teachers to answer checklists to identify hearing impairment is not a useful procedure. Based on these findings, it is not recommended to use solely checklists answered by teachers to screen for the presence of hearing impairment in school-going children.

Findings of the Ling 6-sound Screening Test

General Outcome of the Ling 6-sound Screening Test: From the data of 154 children (308 ears) that were analyzed, 58 children (104 ears) were referred for further testing based on the scores obtained on the screening test (score < 6). These 58 children did not correctly respond to one or more speech sounds of the test. Out of the 58 children, 12 of them did not pass the screening test in only one of their ears and 46 of them did not pass the test in both their ears.

The data were also analyzed for each speech sound separately. This was done to see if the referral differed depending on specific speech sounds. Table 1 summarizes the number of ears that were referred based on the responses for the different speech sounds of the Ling's test.

Table 1: Number of ears referred based on the responses for each of the speech sounds of the Ling's 6-sound test

Speech Sound	Number of ears referred (Total: 308)
/a/	50 (16.23 %)
/i/	66 (21.42 %)
/u/	60 (19.48 %)
/s/	94 (30.52 %)
/sh/	81 (26.29 %)
/m/	86 (27.92 %)

As evident from Table 1, the maximum number of ears was referred for the sound /s/ followed by /m/ and /sh/. McNemar test for related samples was done to check if there was a statistically significant difference in the referral for the six different speech sounds. Statistically significant differences at 0.001 level were found between all combinations of the sounds except for the combinations of /i/ - /u/ and /s/ - /m/. These combinations were not statistically significant even at the 0.05 level.

The results regarding the referral rate based on the Ling's sound test indicated that the perceptual difficulties of these speech sounds were distinctly different. This probably occurred due to the different acoustic characteristics of the speech sounds which tapped perceptual difficulties across different frequencies. The speech sounds /i/ - /u/ and /s/ - /m/ yielded similar results probably because these combinations were similarly affected though they had contrastive frequency responses. An examination of the raw data indicated that a large number of ears had problems in perceiving both the combination of speech sounds. Fifty-four of the 104 ears had difficulty in perceiving both /i/ and /u/ while 80 of the 104 ears had problem perceiving both /s/ and /m/. This highlights that though these speech sounds that had contrastive acoustic patterns they affected the children in a similar manner.

It was also found that /s/ ($p < 0.001$) and /sh/ ($p < 0.001$) were statistically significantly different from all or most of the other Ling's sounds, except the former which did not differ from /m/. It was also seen that these sounds lead to higher referral rates when compared to other speech sounds. This indicates that these two speech sounds were more sensitive to the auditory perceptual problems of children.

The possibility that the perception of the sound could have been due to the presence of low frequency ambient noise in the environment was ruled out. Had the ambient noise affected the test results, it should have affected all the children in a similar manner. Further, the biological calibration that was carried out regularly ensured that the ambient noise did not affect the results. Thus, it can be inferred that the test tapped the auditory perceptual difficulties of the children and was not affected by extraneous factors such as ambient noise.

*Sensitivity and Specificity of the Ling's 6-sound Screening Test:*In order to find out the effectiveness of the Ling's 6-sound screening test in identifying hearing loss, the findings of it were compared with that of the standard tests that included pure-tone tests and standard screening tests (tympanometry, acoustic reflexes & DPOAE). The sensitivity and specificity were determined using a similar procedure as was done for the screening checklist. Sensitivity of the Ling's 6-sound (Table 2) test was found to be highest (82 %) when it was compared with the gold standard test, DPOAE followed by tympanometry (77.6%). Almost similar sensitivity was found when it was compared to PTA 1 (76.7%%), PTA 2 (68.9%) and acoustic reflex test (72.5%). It was the least (62.9%) for the combination of PTA and tympanometry.

The specificity for the Ling's 6-sound test (Table 2) did not vary much when it was compared with different standard procedures. The highest specificity was found when compared with the acoustic reflex test (90.3 %) followed by that of a combination of PTA and DPOAE (89.1 %). Comparable specificity was seen when other tests were used as the standard (PTA 2, tympanometry, combination of PTA and tympanometry). The specificity was the least when compared with PTA 1. Kappa coefficient revealed that there was a good agreement that ranged from 0.47 to 0.64 between the results of Ling's screening test and the various standard tests. This agreement, in all the instances, was statistically significant ($p < 0.05$).

Predictive Values for the screening test: The positive and the negative predictive values for the Ling's 6-sound screening test were high. This resulted in a low over referral and under referral rate for the screening test with respect to different standard tests. The Ling's 6-sound screening test had a relatively lower over referral (18.3% to 50%) and very low under referral (7.4% to 24%) rates than the hearing screening checklist. It can therefore be concluded that this screening test is a better predictor of hearing loss than the screening checklist that was used in the present study.

The agreement of each of the Ling's 6 speech sounds with the standard tests was also determined using Kappa test of agreement (Figure 1). Among the 6 sounds, /s/ (0.48 to 0.63) followed by /sh/ (0.46 to 0.61) had the highest agreement with acoustic reflex test and DPOAE. However, these levels of agreements were generally lower than the values got when the scores of all the 6 sounds (0.49 to 0.64) were combined and used. This probably happened since the 6 different speech sounds covered a range of frequencies that were able to detect hearing losses that occurred in different frequency regions. In contrast, isolated speech sounds probably

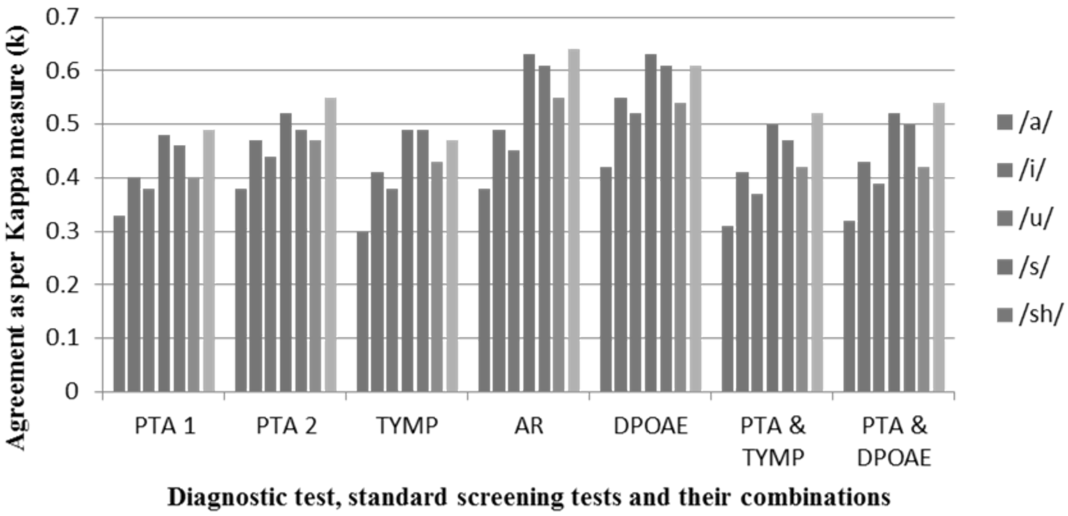


Figure 1: Agreement for each sound with different diagnostic and standard screening test.

Table 2: Sensitivity and specificity of the checklist and the screening test combined and in isolation

		Diagnostic Tests		Standard screening tests		
		PTA 1	PTA 2	TYMP	AR	DPOAE
Checklist & screening test combined	Sensitivity	69.2 %	76.1 %	85 %	85.9 %	86.7 %
	Specificity	66 %	67.8 %	63 %	68.5 %	66.7 %
Checklist	Sensitivity	38.5 %	36.7 %	49.3 %	42.4 %	42.2 %
	Specificity	72.8 %	74.9 %	76.3 %	76.4 %	75.6 %
Ling's screening test	Sensitivity	76.7 %	68.9 %	77.6 %	72.5 %	82 %
	Specificity	79.6 %	85.4 %	78.4 %	90.3 %	84 %

Note: PTA 1 = pure tone average of 500 Hz, 1 kHz and 2 kHz; PTA 2 = pure tone average of 1 kHz, 2 kHz and 4 kHz; TYMP = tympanometry; AR = acoustic reflexes

tapped only specific frequencies and miss identifying hearing problems in other frequencies. Hence, the use of isolated speech sounds to speed up the screening process is not recommended. Using all the 6 sounds is suggested to improve the sensitivity, specificity as well as reduce the over referral and under referral.

The findings of the Ling's 6-sound screening test indicate that a fairly large number of children were suspected to have a hearing loss (Table 1). These findings are supported by previous studies done on prevalence of hearing loss among school children. Mishra, Bhatia and Bhatia (1961) investigated 1390 school going children in Lucknow city of North India. They found that the prevalence of hearing loss was 34%, with the majority of children having conductive hearing loss. Kapur (1965) found the prevalence of hearing loss among children aged between 5 to 15 years to be 18.6%. It was opined that this high prevalence was because of the presence of middle ear pathology. Recently, Sarafraz and Ahmadi (2009) found that out of a total of 785 school going children studying in first and second grade, 306 (39%) had hearing loss. This finding was based on tympanometry test results which substantiated the high prevalence of middle ear disorders in school age children.

Thus, it is possible that the high referral rate seen in the Ling's 6-sound screening test in the present study was on account of the presence of a middle ear problem which is commonly seen in school children. The sensitivity and specificity of the Ling's 6-sound screening test confirms that the test is a valid and useful procedure to detect hearing loss in school going children.

Findings of the Combination of Hearing Checklist & Ling 6-sound Screening Test

General outcome of the screening checklist & screening test when used together: In order to find out the effectiveness of the combined findings of the screening checklist and the screening test, the joint findings were compared with the standard test findings. The sensitivity and specificity were calculated and tabulated from the decision matrix for the combined screening protocol. The decision of 'pass' was made only if the children 'passed' both the checklist and the screening test. Likewise, the decision of 'refer' was made if the children were 'referred' based on any one or both the screening procedures.

Of the 308 ears, 160 passed both the checklist and the Ling's 6-sound test, whereas 46 ears were referred from both of them. The number of ears that were negative on the checklist but positive on the screening test was 44 while 58 ears that were positive on the screening test, had negative results on the checklist.

Sensitivity and specificity of the screening checklist and the screening test when used together: Table 2 summarizes the sensitivity and specificity of the screening procedures when the outcome of the Ling's 6-sound screening test and the checklist were combined and in isolation. It can be observed in Table 2 that the sensitivity of the combined screening procedures (checklist & Ling's 6-sound screening test) was marginally better or comparable to the Ling's 6-sound test in isolation. However, the specificity dropped when the combination was used. Thus, it is recommended that the Ling's 6-sound screening test could be used in isolation in order to have a balanced sensitivity and specificity. Further, the checklist, when used in isolation, had a very low sensitivity and hence is not recommended to be used separately.

The findings of the present study is in consonance with that of by Curry (1950), who opined that the task of identifying hearing problem should not be carried out on the basis of teachers' referrals. From the present study, it is recommended that the Ling's 6 sound screening test be used in isolation instead of along with a checklist for school screening programs. This screening test can be carried out without expensive audiological equipment, provided adequate measures are taken to calibrate the output signal from the computer that is required to be used.

Test-retest reliability measures

Reliability of the screening checklist was determined by administering the alpha reliability co-efficient test. Since there was no change in the responses of the teachers, the alpha reliability co-efficient was 1. This high reliability probably occurred since the teachers answered the checklist for the second time within a span of just three days. However, these results confirm that the teachers were able to provide consistent and reliable responses.

The test-retest reliability of the Ling's 6-sound screening test that was checked on 16 children was found to be high. This was established since the alpha reliability co-efficient was greater than 0.6 [? = 0.73, ($p < 0.05$)].

The testretest Reliability of the standard audiological diagnostic tests (pure-tone audiometry, tympanometry, acoustic reflexes and DPOAE) indicated that there was no change in the responses after 3 days. This confirmed the reliability of the responses obtained from the participants.

Conclusions

In general, a single symptom of the screening checklist was not a good indicator of the presence of hearing loss. Each of the questions was found to be unique because they tapped different aspects related to hearing. Hence, it is recommended that, if the checklist is used, all the questions should be utilized as they tapped different aspects related to hearing. The screening checklist 'Signs and symptoms of hearing loss' was not effective in detecting the presence of hearing impairment. The recorded Ling 6-sound screening test was a better predictor of the presence of hearing loss compared to the screening checklist as it had a higher sensitivity and specificity. The screening test also had relatively low over referral and under referral rates as compared to that of the checklist. The combination of the checklist and the screening test did not give rise to significant better results. The sensitivity of the combined protocol ranged from 69.2 % to 86.7 % and specificity ranged from 63 % to 68.5 %, which is almost similar to that of when the Ling's 6-sound screening test alone. Since the inclusion of the checklist did not improve the effectiveness of the screening program, use of the Ling's 6 sound screening test in isolation is recommended. The test is also recommended since it does not require the use of sophisticated instrumentation.

References

Abusaleh, S. (1999). India Human Development Report: A profile of Indian states in the 1990s. National Council of Applied Economic Research. Oxford University Press, New Delhi.

Alpiner, J. G. (1976). Speech language assessment and intervention with school-age hearing impaired children. In Alpiner, J. G. (1976): *Rehabilitative Audiology for Children and Adults.* (pp.188-240). Baltimore: Williams and Wilkins.

American National Standards Institute. (1991). *American National Standard maximum permissible ambient noise levels for Audiometric Test room.* (ANSI S3.1-1991). New York: American National Standards Institute.

American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening: Panel on audiologic assessment.* Rockville, MD: American Speech-Language-Hearing Association.

Bess, F. H., Dodd-Murphy, J., & Parker, R. (1998). Children with Minimal Sensorineural Hearing Loss- Prevalence, Educational performance & Functional Status. *Ear and Hearing, 19,* 339-354.

Curry, E. T. (1950). The Efficiency of Teacher Referrals in a School Hearing Testing Program. *Journal of Speech and Hearing Disorders, 15*, 211-

214.

FitzZaland, R. E., & Zink, G. D. (1984). A Comparative Study of Hearing Screening Procedures. *Ear and Hearing, 5*(4), 205-210.

Glorig, A., & House, H. P. (1957). A new concept of auditory screening. *American Medical Association: Archives of Otolaryngology, 66,* 228-232.

Harford, E. R., Bess, F. H., Bluestone, C. D. & Klein, J. *D.* (1978). Use of acoustic Impedance Measurement in Screening for Middle Ear Disease in Children. In Harford, E., Bess, F., Bluestone, C., & Klein, J. (Eds.). *Impedance Screening for Middle Ear Disease in Children.* (pp. 230). New York: Grunne & Stratton, Inc.

Kapur, Y. P. (1965). A Study of Hearing Loss in School Children In India. *Journal of Speech and Hearing Disorders, 30,* 225-233.

Kumar, S., & D'Mello, J. (2006). Identifying Children At-Risk for Speech and Hearing Disorders- A Preliminary Survey Report from Hyderabad, India. *Asia Pacific Disability Rehabilitation Journal, 17*(2), 101-108.

Krueger, W. W. O., & Ferguson, L. (2002). A comparison of screening methods in school-aged children. *Otolaryngology- Head and Neck Surgery, 127*(6), 516-519.

Lewis, N., Dugdale, A., Canty, A., & Jerger, J. (1975). Open Ended Tympanometric Screening. *American Medical Association: Archives of Otolaryngology, 101,* 722-725.

Lo, P., Tong, M., Wong, E., & Hasselt, C. V. (2006). Parental suspicion of hearing loss in children with otitis media with effusion. *European Journal of Pediatrics, 165*(12), 851-857.

Martin, F. N. (1991). The pediatric patient. In *Introduction to Audiology. (*pp. 395). New Jersy: Prentice-Hall Inc.

Maxon, A. B., & Brackett, D. (1981). Mainstreaming Hearing Impaired Children. In Bradford, L. J., & Martin, F. N. (Eds.). *Audiology- An Audio Journal for Continuing Education.* 6. New York: Grune and Stratton Inc.

Mencher, G. T., & McCulloch, B. F. (1970). Auditory screening of kindergarten children using VASC. *Journal of Speech and Hearing Disorders, 35,* 241-247.

Mishra, R. N., Bhatia, M. L., & Bhatia, B. P. R. (1961). Investigations of Hearing in School Children. *Indian Journal of Otolaryngology, XIII,* 107-127.

National Sample Survey Organisation. (1991). A report on disabled persons. (NSSO, 1991). New Delhi: National Sample Survey Organisation.

National Sample Survey Organisation. (2002). A report on disabled persons. (NSSO, 2002). New Delhi: National Sample Survey Organisation.

Nikam, S., & Dharamraj. (1971). School screening programme in Mysore city. *Journal of All India Institute of Speech and Hearing, 1,* 28-32.

Niskar, A. S., Kieszak, S. M., Holmes, A., Esteban, E., Rubin, C., & Brody, D. J. (1998). Prevalence of hearing loss among children 6 to 19 years of age: the Third National Health and Nutrition Examination Survey. *Journal of American Medical Association, 279*(14), 1071-1075.

Nix, G. W. (1977). The least restrictive environment. *The Volta Review, 79,* 287-296.

Northcott, W. H. (1972). Hearing Impaired Pupil in the Classroom. *The Volta Review, 74,* 105-108.

Northern, J. L., & Downs, M. D. Eds. (2002). *Hearing in Children* (5^{th} Ed), 291. Lippincott: Williams & Wilkins.

Norton, M. C., & Lux, E. (1960). Double Frequency Auditory Screening in Public Schools. *Journal of Speech and Hearing Disorder, 25,* 293-299.

Nozza, R., Sabo, D., & Mandel, E. (1997). A role for otoacoustic emissions in screening for hearing impairment and middle ear disorders in school age children. *Ear and Hearing, 18*(3), 227-239.

Olusanya, B. (2001). Early detection of Hearing Impairment in a Developing Country: What Options? *Audiology, 40,* 141-147.

Paradise, J. L., & Smith, C. (1979). Impedance screening for preschool children: State of the art. In Harford, E., Bess, F., Bluestone, C., & Klein, J. (Eds.). *Impedance Screening for Middle Ear Disease in Children.* (pp. 113-122). New York: Grune & Stratton, Inc.

Persons with Disabilities (*Equal opportunities, Protection of Right and Full Participation*) Act-1995. Ministry of Law, Justice and Company affairs. Government of India. New Delhi.

Ritchie, B., & Merklein, R. (1972). An evaluation of the efficiency of the verbal auditory screening for children (VASC). *Journal of Speech and Hearing Research, 15,* 280-286.

Sarafraz, M., & Ahmadi, K. (2009). A practical screening model for hearing loss in Iranian school-aged children. *World Journal of Pediatrics, 5*(1), 46-50.

Yoshinaga-Itano, C. (2003). From screening to early identification and intervention: discovering predictors to successful outcomes for children with significant hearing loss. *Journal of Deaf Studies and Deaf Education, 8*(1), 11-30.

Yoshinaga-Itano, C., Sedey, A., Coulter, D., & Mehl, A. (1998). Language of early and later identified children with hearing loss. *Pediatrics, 102*(5), 1161-1171.

Relationship between Auditory Temporal Processing and Working Memory

[1]Rishitha Umesh Hosabettu & [2]Ajith Kumar U.

Abstract

The study evaluated the effect of aging on auditory temporal processing, speech perception abilities and working memory and assessed the relationship among them. A total of 30 young adults in the age range of 18-30 years and 30 geriatric individuals in the age range of 60 to 70 years with normal hearing sensitivity participated in the study. The study was divided into 3 experiments- Psychoacoustic experiments, Speech perception experiment and working memory measures. Psychoacoustic experiments included temporal processing measures- gap detection thresholds, modulation detection threshold for sinusoidally amplitude modulated noise and duration pattern scores. Speech perception experiment involved assessing speech perception scores for sentences at 20 dB, 15 dB, 10 dB, 5 dB, 0 dB, -5 dB, -10 dB signal to noise ratios. Working memory measures contained digit forward, digit backward and operation span test. The results revealed that the gap detection thresholds and duration pattern scores declined with age whereas, aging did not show an effect on modulation detection thresholds. All the working memory measures digit forward, digit backward and operation span task showed deterioration with age. Speech perception in noise in the geriatric group was comparable to that of adults at favourable SNRs (+20, +15, +10, +5 dB SNR) but as the SNR became poorer (0, -5, -10 dB SNR) the geriatric group had significant deterioration when compared to adults. Thus, it can be concluded that working memory has a significant influence and relationship with the temporal processing and speech perception in noise.

Keywords: *Aging, temporal processing, speech perception, working memory.*

Introduction

The auditory system analyses sound signal in three basic domains- frequency, intensity and time. Time is an important domain in hearing since most of the sounds fluctuate over time. The perception of the temporal characteristics of a sound or the alteration of durational characteristics within a restricted or defined time interval is called temporal processing (Musiek et al., 2005).

Temporal processing abilities are known to be of crucial importance in daily listening environment. Perception of temporal parameter of sound is important for a wide range of auditory behaviours including rhythm perception, periodicity pitch discrimination, duration discrimination and phoneme discrimination. Furthermore, temporal processing plays a crucial role in language comprehension, perception of prosodic distinctions and speech perception in ambiguous conditions (Chermak & Musiek, 1997). Speech perception becomes poorer in the presence of noise since the presence of noise reduces the temporal variation of the waveform by filling the valleys of the amplitude spectrum which leads to ambiguity in speech. Timing approximation requires some amount of cognitive skills too (Gooch, Stern & Rakitin, 2009). Some researches indicate the associations among working memory, timing, and aging (Brown, Vousden & McCormack, 1999; Baudouin, Vanneste, Pouthas & Isingrini, 2006).

Working memory enables an individual to temporarily store the information and manipulate it if necessary. Broadway and Engle (2011) reported that individuals with low working memory capacity were less sensitive compared to individuals with high working memory in temporal discrimination tasks. Functional magnetic resonance imaging experiments have revealed prefrontal cortex activation when retrieving temporal context information (Rajah, Ames & D'Esposito, 2008). Prefrontal cortex also controls the working memory (Kane & Engle, 2002). Thus, both the temporal processing and working memory skills share a common anatomical site. Hence, it can be hypothesized that temporal processing abilities depend on cognitive functions such as working memory of the individual.

Aging is a natural process which affects all the systems of the body including the auditory system. Age related changes occur anatomically and physiologically as well as peripherally and centrally. Psychophysical evidence documents a broad decline in a variety of auditory abilities because of chronological aging (Zec, 1995). The geriatric group appears to have poorer frequency discrimination compared to adults. Geriatrics with normal hearing thresholds exhibit larger intensity discrimination thresholds with the largest age related changes occurring for the low frequency tones (Murphy, Bruce, Filippo & Giampaolo, 2006). Hence, aging causes auditory processing deficits. Thus, deterioration in temporal processing is not unexpected.

Parra, Iorio, Mizahi and Baraldi (2004) reported that the elderly individuals with normal hearing have temporal patterning ability less than young subjects with normal

[1]Email: rishithahosabettu89@gmail.com
[2]Reader of Audiology, Email: ajithkumar18@gmail.com

hearing. Kumar and Sangamnatha (2011) extensively studied gap detection thresholds, duration discrimination, modulation detection thresholds and duration pattern scores across different age groups spanning from 20 years to 85 years. They stated that there was deterioration in scores in all the temporal processing skills as age advanced. The maximum decline was observed in the 60 years and above age group. Daniels (2011) used electrophysiological measures to assess gap detection thresholds in adults and geriatrics. The geriatric group showed delayed P2 latency compared to the young adults. The geriatric group also had an overall poor wave morphology compared to adults.

Aging causes an overall decline which also includes the working memory abilities. Age related decrements are found in working memory tasks (Light & Anderson, 1985; Spilich, 1983; Wright, 1981). The decline in the working memory is evident when the complexity of the task is increased. There is an increase in the time required to respond by the geriatrics as compared to the adults as the grammatical complexity of the sentence was increased (Gick, Craik & Morris, 1988; Baddeley & Hitch, 1974).

Supporting evidences for the decline in temporal processing and working memory with the age also comes from speech perception studies that have used complex and acoustically degraded speech stimulus. It has been reported that geriatrics experience increased difficulty in understanding speech in noise (Cooper & Gates, 1991). This difficulty in perception may be because of the reduced temporal information received by the listener due to the noise (Tremblay, Piskosz & Souza, 2003). Speech perception in the presence of noise also requires memory (Zacks, Hasher & Li, 2000) since it demands the ability to filter out irrelevant competing noise (Tun & Wingfield, 1999; Tun, O'Kane & Wingfield, 2002).

Several studies have demonstrated that temporal processing and speech perception abilities decline with age even when the hearing thresholds are within normal limits (Kumar & Sangamnatha, 2011; Gordon-Salant & Fitzgibbons, 1995; Cruickshanks et al., 1998). One of the factors that influence speech perception and temporal processing abilities is the working memory (Broadway & Engle, 2011; Wong et al., 2009). Age-related decline in speech perception in noise may be supplemented by increased usage of general cognitive abilities like working memory and attention as a means of compensation for these declines (Wong et al., 2009). Therefore, geriatrics who experience decline in memory or attention are particularly affected by decrease in speech perception (Shinn-Cunningham & Best 2008). Hence, the present study was taken up to assess the possible effect of aging on temporal processing, working memory and speech perception in noise and the relationship among these dependent variables.

Therefore, the aim of the study was to assess the effect of aging on auditory temporal processing, speech perception abilities and working memory and assess the relationship between them.

Method

Participants

A total of 60 participants contributed to the present research. The participants were divided into 2 groups. Group I consisted of 30 young adults in the age range of 18 to30 years. The Group II consisted of 30 normal hearing geriatric individuals in the age range of 60 to70 years. Normal hearing sensitivity was operationally defined as audiometric thresholds within 15 dB HL in octave frequencies from 250 Hz to 2 kHz and thresholds within 30 dB HL at 4 kHz and 8 kHz. A brief case history was noted before initiating the study. The participants with history of middle ear pathology or surgery and complaint of any neurological problems were not included in the study.

A modified version of the Hughson-Westlake procedure (Carhart & Jerger, 1959) was used to measure the hearing thresholds of all participants using a calibrated clinical audiometer (Maico MA52) in an acoustically treated booth with ambient noise level within permissible limits (ANSI, 1999). All participants in the group I had air and bone conduction hearing thresholds less than 15 dB HL at the octave frequencies between 250 Hz and 8 kHz. 9 out of the 30 participants in group II had hearing thresholds up to 30 dB HL at 4 kHz and 8 kHz and at other frequencies the thresholds were within 15 dB HL. The study was divided into 3 experiments- Psychoacoustic experiments, speech perception experiment and working memory measures.

Psychoacoustic Experiments

Stimulus and Procedure: All of the temporal processing measures except for the duration pattern were carried out using 'mlp' tool box (Grassi & Soranzo, 2009) which implements maximum likelihood procedure in Matlab. The maximum likelihood procedure employs a large number of candidate psychometric functions and after each trial calculates the probability (or likelihood) of obtaining the listeners response to all of the stimuli that have been presented given each psychometric function. The psychometric function yielding the highest probability is then used to determine the stimulus to be presented on the next trial. Within about 12 trials, the maximum likelihood procedure usually converges on a reasonably stable estimate of the most likely psychometric function, which then can be used to estimates the threshold (Green, 1990: 1993). Stimuli were generated at 44,100 Hz sampling rate. A two-interval alternate force choice method using a 'maximum likelihood procedure' was employed to track an 80% correct response criterion. Thirty test trails were used. During each trial, stimuli were presented in each of two intervals; one interval contained a reference stimulus, the other interval the variable stimulus. The participant indicated which interval contained the variable stimulus after each trial.

*Gap Detection Thresholds:*The participant's ability to detect a temporal gap in the centre of a 750 ms broadband noise was measured. The noise had 0.5 ms cosine ramps at the beginning and end of the gap. In a two-interval alternate forced-choice task, the standard stimulus was always a 750 ms broadband noise with no gap whereas the variable stimulus contained the gap.

*Modulation Detection Thresholds:*Temporal modulation refers to a reoccurring change (in frequency or amplitude) in a signal over time. A 500 ms Gaussian noise was sinusoidally amplitude modulated at modulation frequencies of 8 Hz, 20 Hz, 60 Hz and at 200 Hz. Noises had two 10 ms raised cosine ramps at the onset and offset. Subject had to detect the modulation and tell which interval had the modulated noise. Modulated and un-modulated stimuli were equated for total root mean square (rms) power. Depth of the modulated signal was varied according to the participant's response up to an 80% criterion level. The modulation detection thresholds were expressed in dB by using the following relationship:

Modulation detection thresholds in dB = $20 \times log_{10}\, m$

Where m= modulation detection threshold in percentage.

*Duration Pattern Scores:*The duration pattern was administered in the manner described by Musiek, Baran and Pinheiro (1990). A 1000 Hz pure tone was generated at 44,100 sampling frequency with two different durations (i.e. short 250 ms and long 500 ms), using Audacity software (ver. 1.3). By combining these two durations in three-tone patterns six different patterns were generated (Short, Short Long, Short Long Short, Long Long Short, Long Short Short, Short Long Long, Long, Short Long). Inter-stimulus interval was 250 ms within a tone sequence and 6 seconds between two tone sequences. Following practice trails, 30 test items were administered. Participants were asked to verbally repeat the sequence.

Speech Perception Experiment

Speech perception in noise was evaluated using the test developed by Methi, Avinash and Kumar (2009). Seven equivalent lists from the original test were selected for the present study. Each list contained 7 sentences mixed with the eight talker speech babble noise at different signal to noise ratios (SNRs). First sentence in each list was at +20 dB SNR, second sentence was at +15 dB SNR, third sentence was at +10 dB SNR, fourth sentence was at +5 dB SNR, fifth sentence was at 0 dB SNR, sixth sentence was at -5 dB SNR and last sentence was at -10 dB SNR. Each sentence had 5 key words. These sentences were presented through a personal computer (Dell Inspiron 15R) at comfortable listening levels through circumaural headphones (Intex). The listener's task was to repeat the sentences presented and each correctly repeated key word was awarded one point for a total possible score of 35 points per list.

Working Memory Measures

*Auditory Digit Span:*Auditory working memory was assessed using the auditory digit span. The auditory digit span is divided into forward and backward phase. The numbers were recorded from 1 to 9 and 6 lists were prepared with increasing level of difficulty with level 1 being the easiest and level 6 being the toughest. Level 1 contained 3 digits while the level 6 contained 8 digits which were randomly presented. An inter stimulus interval of 25 ms was maintained for all the levels. These clusters of digits were presented and the participants were asked to repeat the numbers in same or backward order for digit forward and digit backward task respectively. The scoring was based on the number of digits correctly repeated by the participant.

Operation Span Task (OST): The procedure and scoring was adapted from versions of the OST used by Kane et al. (2004). In the OST, each element consisted of a mathematical operation and a word (e.g., 3+5-4=4, yes or no? /mara/). The words used in the test were familiarity rated initially and then the most familiar and least familiar words were eliminated from the list. The participant's task was to read the math problem aloud, say " yes" or " no" to indicate whether the given answer is correct or incorrect and then say the word. After all the elements in an item are presented, the participants were required to write the words in correct serial order. The difficulties of the items were randomized such that the numbers of elements were unpredictable at the outset of an item. Guidelines recommended by Conway, Cowan, and Bunting (2001) were followed during the scoring. A score of 1 was assigned for every word correctly recalled which sums up to a maximum score of 20.

Statistical Analysis

Descriptive statistics was computed to calculate the mean and standard deviation for the temporal processing measures and speech in noise test across the two groups. Analysis of covariance (ANCOVA) was administered to assess the effect of aging on gap detection threshold and duration pattern scores. Multivariate analysis of covariance (MANCOVA) was administered to assess the effect of aging on modulation detection thresholds for sinusoidally amplitude-modulated noise and speech perception in noise by eliminating the influence of working memory and minimal hearing loss.

Independent t test was computed to assess the effect of age on working memory measures. Karl Pearson's coefficient correlation was calculated to assess the correlation between temporal processing and working memory, temporal processing and speech perception in noise.

Results

Appropriate statistical analysis was computed using SPSS version 20. The following statistical procedures were used to analyse the data.

Effect of Age on Temporal Processing

Gap detection threshold (GDT): Figure 1 shows the mean GDT along with the one standard deviation (SD) variation for the adult and the geriatric group. The mean scores noticeably indicate that the performance of the adult group was better when compared to the geriatric group. Additionally, the variability as evidenced by the standard deviations was more for the geriatric group when compared to the adult group. ANCOVA was performed to assess the significance of differences between the mean GDT between two groups. As working memory and hearing thresholds can affect the GDT, these were used as co-variates (numerical independent variables) in the model. ANCOVA results showed a significant main effect of subject group on GDT [$F(1, 54) = 15.461$ $p<0.05$] after controlling the effect of minimal hearing loss in the high frequency region (4 kHz & 8 kHz) and working memory. The covariate OST significantly influenced the participant's GDT [$F(1, 54) = 15.879$ $p<0.05$]. However, the hearing thresholds [$F(1, 54) = 0.410$ $p>0.05$], digit forward [$F(1, 54) = 3.228$ $p>0.05$] and digit backward [$F(1, 54) = 1.811$ $p>0.05$] did not influence the GDT of the participants.

Modulation detection threshold (MDT): Figure 2 shows the mean for MDT at 8 Hz, 20 Hz, 60 and 200 Hz along with the one SD variation for the adult and the geriatric group. From the Figure 2 it can be seen that mean modulation detection thresholds were better in the adult group as compared to the geriatric group. Additionally, the variability as evidenced by the standard deviations was more for the geriatric group when compared to the adult group. MANCOVA was performed with MDT at 8 Hz, 20 Hz, 60 Hz and 200 Hz as dependent variable, subject group as independent variable and average of hearing thresholds in high frequencies (2 kHz, 4 kHz and 8 kHz in both the ears) and working memory measures as covariate. MANCOVA results showed no significant main effect of subject group on MDT 8 Hz [$F(1, 54) = 0.877$ $p>0.05$], MDT 20 Hz [$F(1, 54) = 2.412$ $p>0.05$], MDT 60 Hz [$F(1, 54) = 4.592$ $p>0.05$] and MDT 200 Hz [$F(1, 54) = 0.156$ $p>0.05$] after factoring out the effect of minimal hearing loss and working memory. This means that modulation detection thresholds were comparable between the adults and geriatrics at all the modulation frequencies tested.

Duration pattern scores: Figure 3 shows the mean duration pattern scores along with the one SD variation for the adult and the geriatric group. The Figure 3 illustrates that the mean duration pattern scores for adults was much higher than the geriatric group. Additionally, the variability as evidenced by the standard deviations was more for the geriatric group when compared to the adult group. ANCOVA was performed with duration pattern scores as dependent variable, age as independent variable and average of hearing thresholds in high frequencies (2 kHz, 4 kHz and 8 kHz in both the ears) and working memory measures as covariate. ANCOVA results showed a significant main effect of subject group on duration pattern scores [$F(1, 54) = 9.192$ $p<0.05$] after factoring out the effect of minimal hearing loss and working memory. The covariates hearing thresholds [$F(1, 54) = 5.004$ $p<0.05$], operation span [$F(1, 54) = 4.392$ $p<0.05$] and digit forward [$F(1, 54) = 5.610$ $p<0.05$] significantly influenced the participant's duration pattern scores. However, the digit backward [$F(1, 54) = 0.268$ $p>0.05$] did not influence the duration pattern scores of the participants.

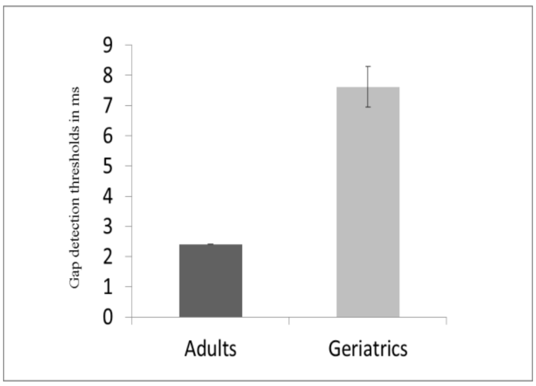


Figure 1: The mean gap detection thresholds in adults and geriatrics. The error bars indicate 1 SD of error.

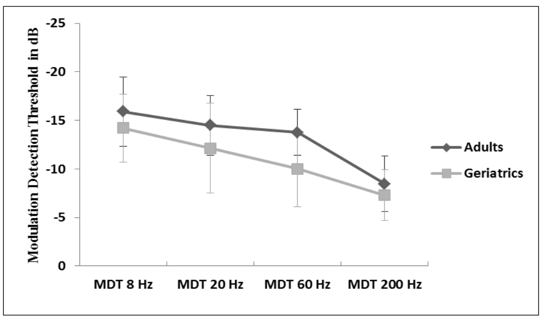


Figure 2: The mean modulation detection thresholds at 8 Hz, 20 Hz, 60 Hz and 200 Hz in adults and geriatrics. The error bars indicate 1 SD of error. [MDT- modulation detection threshold]

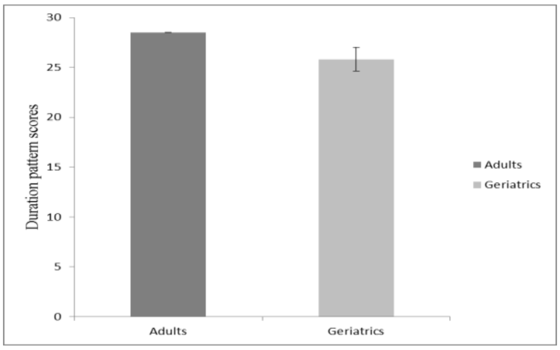


Figure 3: The mean duration pattern scores in adults and geriatrics. The error bars indicate 1 SD of error.

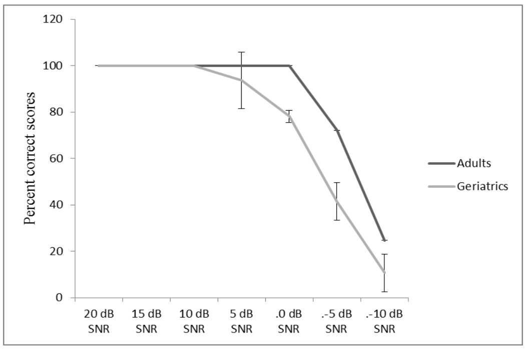


Figure 6: The mean speech in noise scores at 20 dB, 15 dB, 10 dB, 5 dB, 0 dB, -5 dB and -10 dB SNR in adults and geriatrics. The error bars indicate 1 SD of error.

Effect of Age on Working Memory Measures

Figure 4 shows the mean scores for digit forward and digit backward and Figure 5 shows the mean scores for OST along with the one standard deviation (SD) variation for the adult and the geriatric group. The mean scores indicate that the working memory is better for the adult group as compared to the geriatric group. The results of the independent samples t-test revealed that the adult group had significantly better digit forward ($t = 4.175$, $p<0.05$), digit backward ($t = 3.971$, $p<0.05$) and operation span ($t = 4.953$, $p<0.05$) scores when compared to the geriatric group.

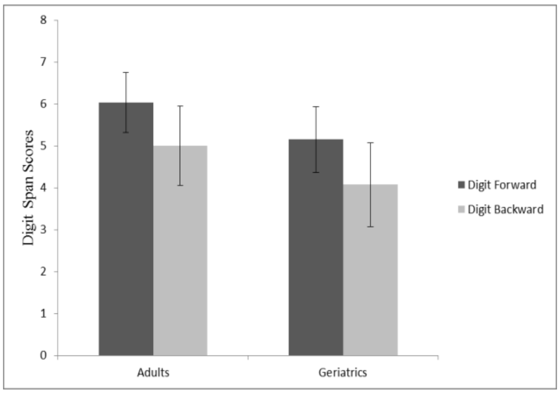


Figure 4: The mean digit forward and digit backward scores in adults and geriatrics. The error bars indicate 1 SD of error.

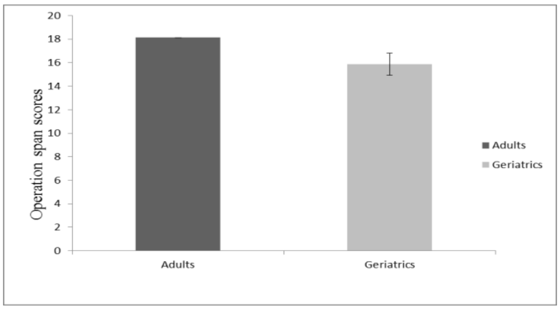


Figure 5: The mean operation span scores in adults and geriatrics. The error bars indicate 1 SD of error.

Effect of age on speech perception in noise (SIN)

Figure 6 shows the mean scores for SIN along with the SD variation for the adult and the geriatric group. The mean scores indicate that the SIN is better for the adult group when compared to the geriatric group especially at higher SNRs. The raw speech perception scores were converted in rationalized arcsine units (rau). The conversion of raw scores to rau scores was done using the formula by Sherbecoe and Studebaker (2004) which was implemented in MATLAB. All the further statistical analysis was carried out using the rau speech perception scores. At +20 dB SNR, +15 dB SNR, +10 dB SNR participants in both the groups obtained 100% correct identification and hence these SNRs were excluded from further statistical analysis. MANCOVA was performed to see the significance of differences in the speech perception scores between the groups. The speech identification scores at 5 dB, 0 dB, -5 dB and -10 dB SNR as dependent variable, subject groups as independent variable and average of hearing thresholds in high frequencies (2 kHz, 4 kHz and 8 kHz in both the ears) and working memory measures were used as covariates in the model. MANCOVA results revealed a significant main effect of subject group on speech perception at 5 dB SNR [$F(1, 54) = 12.79$, $p< 0.05$], 0 dB SNR [$F(1, 54) = 37.611$, $p<0.05$], -5 dB SNR [$F(1, 54) = 22.241$, $p<0.05$] and -10 dB SNR [$F(1, 54) = 6.889$, $p< 0.05$].

Relationship between Temporal Processing and Working Memory

Karl Pearson's correlation co-efficient was computed to evaluate the possible relationship between temporal processing and working memory. Each of the temporal processing measures was correlated with the working memory measures. Data from adult and geriatric were pooled in for this purpose. Table 1, shows the correlation co-efficient 'r' between the variables. The analysis showed a significant negative correlation between GDT, MDT at 8, 20, 60, 200 Hz and all the working memory measures. Duration pattern scores showed a high positive correlation with the working memory measures. A negative correlation indicates that GDT and MDT were better in individuals with higher working memory capacity (WMC) as measured using digit forward, backward and OST. A positive correlation indicates that individuals who had higher WMC also had better duration pattern sores. The levels of significances are mentioned for each of the variables in the table below.

Table 1: Correlation between temporal processing and working memory

Temporal processing measures	Working memory measures		
	Digit forward	Digit backward	OST
GDT	-0.600**	-0.563**	-0.734**
DPT	0.683**	0.660**	0.705**
MDT 8 Hz	-0.416**	-0.385**	-0.388**
MDT 20 Hz	-0.248	-0.296*	-0.415**
MDT 60 Hz	-0.549**	-0.491**	-0.478**
MDT 200 Hz	-0.435**	-0.321*	-0.314*

***Correlation is significant at the 0.01 level (2-tailed).*
**Correlation is significant at the 0.05 level (2-tailed).*

[GDT- Gap detection threshold, DPT- Duration pattern scores, MDT 8 Hz- Modulation detection threshold at 8 Hz, MDT 20 Hz- Modulation detection threshold at 20 Hz, MDT 60 Hz- Modulation detection threshold at 60 Hz, MDT 200 Hz- Modulation detection threshold at 200 Hz]

Relationship between Speech Perception in Noise and Working Memory

Karl Pearson's correlation co-efficient was computed to evaluate the possible relationship between working memory and speech in noise. Each of the working memory measures was correlated with the speech in noise scores at +5, 0, -5 and -10 dB SNR. Data from adult and geriatric were pooled in for this purpose. Table 2, shows the correlation co-efficient 'r' between the variables. The analysis showed a significant positive correlation between all the working memory measures and speech in noise at poorer SNRs ie., 0 dB, -5 dB and -10 dB SNRs. Additionally, OST showed a positive correlation with speech in noise even at 5 dB SNR. A positive correlation indicates that individuals who had higher WMC also had better speech in noise scores. The levels of significances are mentioned for each of the variables in the table below.

Table 2: Correlation between working memory and speech in noise

Working memory measures	Speech in noise test (dB SNR)			
	5	0	-5	-10
Digit forward	0.18	0.39**	0.54**	0.59**
Digit Backward	0.17	0.39**	0.55**	0.61**
OST	0.28*	0.51**	0.66**	0.69**

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).
[OST- operation span task]

Discussion

The main aim of this study was to assess effect of aging on temporal processing, working memory and speech perception in noise. This study also explored the relationship between working memory capacity (WMC) and temporal/speech perception skills. Results revealed that temporal processing (except modulation detection thresholds), speech perception and working memory skills declined with the advancing age. Furthermore, the working memory measures were significantly correlated with the temporal processing and speech perception skills.

Effect of Age on Temporal Processing

Gap detection thresholds and duration pattern scores showed a significant deterioration with age. Several studies in the past quote the evidence for deterioration in gap detection thresholds with age (Robin & Royer, 1987; Moore & Glasberg, 1988; Schneider, Pichora-Fuller, Kowalchuk & Lamb, 1994; Snell, 1997; Kumar & Sangamanatha, 2011). Snell (1997) assessed the gap detection thresholds in young adults and geriatrics with normal hearing sensitivity. He reported a poor gap detection threshold in the geriatric group when compared to the adults. Kumar and Sangamnatha (2011) reported gap detection thresholds to be 8 fold greater in individuals above 70 years of age as compared to individuals in 20to 30 age range. Trainor and Trehub (1989) reported temporal sequencing impairment in elderly listeners irrespective of the hearing loss. Several studies have reported that temporal patterning skills decline with age (Kumar & Sangamnatha, 2011; Parra et al., 2004) especially after the 6^{th} decade of life (Kumar & Sangamnatha, 2011).

Results of the present study also revealed that gap detection thresholds were significantly influenced by the participant's operation span skills. Duration pattern scores were significantly affected by digit forward and operation span skills. This means that both of these measures depend on participants' WMC. To our knowledge this is the first report evaluating the relationship between working memory measures and auditory temporal processing skills. However, there are several indirect evidences in the literature which shows that there is a relationship between temporal processing and cognition in general. Unsworth and Engle (2007) stated that individuals differ in their performance in memory tasks such as serial order recall because of the differences in their WMC. Individuals with low WMC are unable to use the temporal contextual cues to the same extent as the individuals with high WMC. Evidence for changes in temporal judgment is reported throughout the lifespan

(McCormack, Brown, Maylor, Darby & Green, 1999; Baudouin, Vanneste, Pouthas et al. 2006) and markedly tends to differ in WMC as well (Brown et al. 1999). Thus, an association exists among WMC, timing and aging (Baudouin, et al. 2006). Conway and Engle (1994) stated that individuals who were categorised as having high WMC based on operation span task scores demonstrated to have better blocking out or are less affected by distracting information. Conway et al. (2001) stated that individuals with low WMC based on operation span task had difficulty in repeating the stimulus as compared to the high WMC individuals in the presence of competing signal. Broadway and Engle (2011) reported low working memory capacity individuals were less sensitive than the high working memory individuals in the temporal discrimination task. They also reported that individual differences in working memory capacity also had individual differences in temporal discrimination. This finding is supported by the theory of individual differences in working memory capacity (Unsworth & Engle, 2007) and theory of short-term memory (Brown, Preece & Hulme, 2000) which propose that recall and recognition depend on discriminating memory.

Modulation detection thresholds did not show a significant difference between the adults and the geriatrics after eliminating the influence of hearing thresholds and working memory measures. The modulation detection thresholds were comparable between the adults and geriatrics at all the modulation frequencies tested. This is in contrast to other studies which have reported an age related decline in the modulation detection thresholds (Kumar & Sangamanatha, 2011; He, Mills, Ahlstrom, & Dubno, 2008). This discrepancy between the present study and the others may be because previous studies have not controlled the effect of minimal hearing loss in the high frequency region, which is often encountered while testing geriatric individuals, and also the WMC. For example, Kumar and Sangamanatha (2011) reported that modulation detection thresholds deteriorated by the 6^{th} decade for lower modulation frequencies (8 Hz and 20 Hz) and by the 4^{th} decade for higher modulations (60 Hz and 200 Hz). But they did not measure the working memory capacities in their participants and decline in the working memory may be one of the contributors for poor modulation detection thresholds seen in their participants. He et al. (2008) also assessed the modulation threshold in adults and geriatrics. Geriatrics up to mild hearing loss at high frequencies was considered in the study. They reported deterioration in modulation thresholds with age. But, the influence of neither hearing loss nor working memory was controlled in the study. Results of the present study are similar to that of Takahashi and Bacon (1992). They showed that even minimal hearing loss had an effect on modulation detection threshold whereas aging did not show much difference in the modulation detection threshold when hearing loss was controlled. In the current study, effects of these two independent numerical variables ie., hearing loss and working memory were factored out as they were used as covariates in the statistical analysis.

Effect of Age on Working Memory

Results revealed that performance of geriatric individuals were significantly poorer than adults on all the working memory measures that were tested. Verhaeghen and Salthouse (1997) assessed the WMC across age. They reported a significant negative correlation between age and cognition and also reported the decline in memory accelerated after 50 years of age. Lustig and Meck (2001) described an age related decline in the memory. Similar results have been documented by Hasher and Zacks (1988); Babcock and Salthouse (1990) wherein they report a decline in the working memory with increasing age. Hasher and Zacks (1988) justify that age related deficit in filtering or supressing irrelevant information lead to excessive load on WMC and thus reduce performance. One possible reason for this decline could be the reduced ability to attend to the stimuli (Lustig & Meck, 2001). This reduced attention having an effect on the working memory is supported by the controlled attention theory of working memory by Engle and Kane (2004). According to this hypothesis, there is a general component of working memory responsible for guiding attention as well as domain specific components responsible for maintenance of task relevant information. Individuals with high WMC have better attention skills and can maximally make use of domain specific skills and strategies to aid maintenance (Colflesh & Conway, 2007).

Effect of Age on Speech in Noise (SIN)

In favourable SNRs (up to +10 dB SNR), performance of the geriatric group was comparable to that of adult group. However, at less favourable SNRs (5 dB and below up to -10 dB SNR) performance of the geriatric group was significantly worse when compared to adult group. It has been reported that geriatric listeners experience increased difficulty in understanding speech in noise (Cooper & Gates, 1991). Kumar and Sangamnatha (2010) reported a decline in the speech in noise scores in spite of having normal audiometric thresholds after 40 years of age which significantly deteriorated further as the age increased. This difficulty in speech perception in noise may be because of the reduced temporal information received by the listener due to the noise (Tremblay et al. 2003). In the unfavourable condition listening is highly effortful. When the listening conditions are unfavourable words cannot be identified on the basis of the signal cues alone. Stored information must be used to achieve the correct identification. Although, the supportive context in the sentence helps in the lexical access, this is cognitively more demanding when compared to the auditory input is less ambiguous as in better SNR conditions. Older listeners had working memory capacity that was significantly less than the young adults. This decline in the working memory capacity of older adults is one of the reasons for observed poor speech perception scores in older adults.

Relationship between Temporal Processing and Working Memory

Correlation analyses showed that there is a significant relationship between the working memory measures and speech in noise. This means that individuals with high WMC which was measured using digit forward, backward and OST also had better temporal processing skills. The working memory measures, digit forward and digit backward tasks tap the auditory memory of the individual and the OST requires the listener to selectively attend to the words to be recalled. Previous studies have reported that abilities to discriminate short intervals depend on differences in attention (Lustig & Meck, 2001; Vanneste & Pouthas, 1999) or memory (Perbal et al., 2005; Rakitin et al., 2006; Rakitin et al., 2005), and sometimes both (Baudouin, Vanneste, Isingrini & Pouthas 2006; Baudouin et al., 2006). Aging causes deterioration of both memory and attention (Park & Hedden, 2001; Reuter-Lorenz & Sylvester, 2005). Hence, a possible relationship exists between gap detection threshold, modulation detection threshold and working memory.

Temporal patterning requires additional cognitive skills like memory as the complexity of the task rises by increasing the length of the stimulus (Fogerty, Humes & Kewley-Port, 2010). The auditory digit span task taps the memory component of cognition and the load on auditory memory increases by increasing the number of digits in the digit span task. Temporal patterning abilities are thus assumed to be better in individuals with better auditory memory. Hence, there is a relationship between working memory and temporal patterning abilities.

Relationship between working memory and speech in noise (SIN)

The results revealed that SIN deteriorated with age and OST had an influence on the SIN scores. Moreover, the SIN scores showed high correlation working memory measures. The influence of working memory on SIN was seen at 0 dB, -5 dB and -10 dB SNR but not at +5 dB SNR. Thus, the results of the present study shows that greater level of cognition is required for perception of speech in noise when the SNRs are poor and not when the speech is well above the noise levels. Wong et al. (2009) reported similar results based on fMRI studies. The results showed reduced activation in the auditory cortex but an increase in working memory and attention-related cortical areas which are the prefrontal and precuneus regions in geriatrics, especially in the poorer SNR condition. Colflesh and Conway (2007) reported that the selective attention supports the notion that individuals with greater WMC are better able to focus attention and avoid distraction. Conway et al. (2001) also reported that working memory is responsible for maintaining activation to relevant information and suppressing the distracting information.

Conclusions

To summarize, the present investigation showed that auditory temporal processing and speech processing abilities were strongly dependent on WMC. Performance of individuals with high WMC on temporal and speech processing tasks was superior to individuals with low WMC. Therefore, observed speech understanding difficulties of older individuals may be due to combined effect of reduced WMC affecting multiple domains of auditory processing.

References

American National Standards Institute (1999). *Maximum permissible ambient noise levels for audiometric rooms.* (ANSI S3.1-1999). NewYork: ANSI.

Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging, 5,* 421-428.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent Advances in Learning and Motivation, vol 8,* (pp. 47-89). New York: Academic Press.

Baudouin, A., Vanneste, S., Isingrini, M., & Pouthas, V. (2006). Differential involvement of internal clock and working memory in the production and reproduction of duration: A study on older adults. *Acta Psychologica, 121,* 285-296.

Baudouin, A., Vanneste, S., Pouthas, V., & Isingrini, M. (2006). Age-related changes in the duration reproduction: Involvement of working memory processes. *Brain and Cognition, 62,* 17-23.

Broadway, J. M., & Engle, R. W. (2011). Lapsed attention to elapsed time? Individual differences in working memory capacity and temporal production. *Acta Psychologica, 137,* 115-126.

Brown, G. D. A., Vousden, J. I., & McCormack, T. (1999). The development of memory for serial order: A temporal-contextual distinctiveness model. *International Journal of Psychology, 34,* 389-402.

Brown, G. D. A., Preece, T., & Hulme, C. (2000) Oscillator-based memory for serial order. *Psy-*

chological Review, 107, 127-181.

Carhart, R., & Jerger J. (1959). Preferred Methods for Clinical Determination of Pure-Tone Thresholds. *Journal of Speech Language and Hearing Research, 24,* 330-345.

Chermak, G. D., & Musiek, F. E. (1997). *Central auditory processing disorders: New perspectives.* San Diego, CA: Singular.

Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin and Review, 14,* 699-703.

Conway, A. R. A., & Engle, R. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123,* 354-373.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8,* 331-335.

Cooper, J. C., & Gates, G. A. (1991). Hearing in the elderly-the Framingham cohort, 1983-1985: Part II: Prevalence of central auditory processing disorders. *Ear and Hearing, 12,* 304-311.

Craik, F. I., & Salthouse, T. A. (2000). *The handbook of aging and cognition (2nd ed).* US: Lawrence Erlbaum Associates Publishers.

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E. K., Klein, R., Mares-Perlman, J. A., & Nondahl, D. M. (1998). Prevalence of hearing loss in older adults in Beaver Dam, WI: The Epidemiology of Hearing Loss Study. *American Journal of Epidemiology, 148,* 879-86.

Daniels, S. B. (2011). *Electrophysiological Measures of Gap Detection Threshold in Younger and Older Normal Hearing Adults.* ProQuest Dissertations and Theses. http://digitalcommons.uconn.edu/dissertations/AAI3475573/

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *The psychology of learning and motivation, 44,* 145-199.

Fogerty, D., Humes, L. E., & Kewley-Port, D. (2010). Auditory temporal-order processing of vowel sequences by young and older adults. *Journal of the Acoustical Society of America, 127,* 2509-2520.

Gick, M. L., Craik, F. I. M., & Morris, R. G. (1988). Task complexity and age differences in working memory. *Memory & cognition, 16,* 353-361.

Gooch, C., Stern, Y., & Rakitin, B. C. (2009). Evidence for age-related changes to temporal attention and memory from the choice time production task. *Aging, Neuropsychology, and Cognition, 16,* 285-310.

Gordon-Salant, S., & Fitzgibbons, P. J. (1995). Recognition of Multiply Degraded Speech by Young and Elderly Listeners. *Journal of Speech and Hearing Research, 38,* 1150-1156.

Grassi, M., & Soranzo, A. (2009). MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods, 41,* 20-28.

Green, D. M. (1990). Stimulus selection in adaptive psychophysical procedures. *Journal of the Acoustical Society of America, 87,* 2662-2674.

Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in yes-no task. *Journal of the Acoustical Society of America, 93,* 2096-2105.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower Ed.,*The Psychology of Learning and Motivation, vol 22*(pp. 193-225). New York: Academic Press.

He N. J., Mills J. H., Ahlstrom J. B., Dubno J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *Journal of the Acoustical Society of America, 124*, 3841-3849.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory, executive attention and general fluid intelligence. *Psychonomic Bulletin and review, 9,* 637-671.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189-217.

Kumar, U. A., & Sangamanatha, A. V. (2011). Temporal processing abilities across different age groups. *Journal of the American Academy of Audiology, 22,* 5-12.

Light, L. L., & Anderson P. A, (1985). Working memory capacity, age and memory for discourse. *Journal of Gerontology, 40,* 737-747.

Lustig, C., & Meck, W. H. (2001). Paying attention to time as one gets older. *Psychological Science, 12,* 478-484.

McCormack, T., Brown, G. D. A., Maylor, E. A., Darby, A., & Green, D. (1999). Developmental changes in time estimation: Comparing childhood and old age. *Developmental Psychology, 35,* 1143-1155.

Methi, R., Avinash, & Kumar, U. A. (2009). Development of sentence material for Quick Speech in Noise test (Quick SIN) in Kannada. *Journal* *of the Indian Speech, Language and Hearing Association, 23,* 59-65.

Moore, B. C. J., & Glasberg, B. R. (1988). Gap detection with sinusoids and noise in normal, impaired and electrically stimulated ears. *Journal of the Acoustical Society of America, 83,* 1093-1101.

Morris, R. G., Gick, M. L., & Craik, F. I. M. (1988). Processing resources and age differences in working memory. *Memory & Cognition, 16,* 362-366.

Murphy, D. R., Bruce, A. S., Filippo, S., & Giampaolo, M. (2006). A Comparison of Higher Order Auditory Processes in Younger and Older Adults. *Psychology and Aging, 21,* 763-773.

Musiek, F. E., Baran, J. A., & Pinheiro, M. L. (1990). Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology, 29,* 304-313.

Musiek, F., Shinn, J., Jirsa, B., Bamiou, D., Baran, J., & Zaidan, E. (2005). GIN (Gaps-In-Noise) test performance in subjects with confirmed central auditory nervous system involvement. *Ear and Hearing, 26,* 608- 618.

Park, D. C., & Hedden, T. (2001). Working memory and aging. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger(Eds.) *Perspectives on Human Memory and Cognitive Aging: Essays in honour of Fergus Craik* (pp. 148-160). East Sussex, UK: Psychology Press.

Parra, V. M., Iorio, M. C., Mizahi, M. M., & Baraldi, G. (2004). Frequency and duration patterns tests in elderly people with normal hearing sensitivity. *Revista Brasileira de Otorhinolaringologia, 70,* 517-523.

Perbal, S., Deweer, B., Pillon, B., Vidailhet, M., Dubois, B., & Pouthas, V. (2005). Effects of internal clock and memory disorders on duration reproductions and duration productions in patients with Parkinson's disease. *Brain and Cognition, 58,* 35-48.

Rajah, M. N., Ames, B., & D'Esposito, M. (2008). Prefrontal contributions domain-general executive control processes during temporal context retrival. *Neuropsychologia, 46,* 1088-1103.

Rakitin, B. C., Scarmeas N., Li, T., Malapani, C., & Stern, Y. (2006). Single-dose levodopa administration and aging independently disrupt time production. *Journal of Cognitive Neuroscience, 18,* 376-387.

Rakitin, B. C., Stern, Y., & Malapani, C. (2005). The effects of aging on time production in delayed free-recall. *Brain and Cognition, 58,* 17-34.

Reuter-Lorenz, P. A., & Sylvester, C.Y. (2005). The cognitive neuroscience of aging and working memory. *The Cognitive Neuroscience of Aging* 186-217. Oxford University Press,

Robin, D. A., & Royer, F. L. (1987). Auditory temporal processing: two-tone flutter fusion and a model of temporal integration. *Journal of the Acoustical Society of America, 82,* 1207-1217.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. *Journal of the Acoustical Society of America, 95,* 980-991.

Sherbecoe, R., & Studebaker, G. (2004). Supplementary Formulas and Tables for Calculating and Interconverting Speech Recognition Scores in Transformed Arcsine Units. *International Journal of Audiology, 43,* 442-448.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification, 12,* 283-299.

Smith, E., E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P., & Koeppe, R. (2001). The neural basis of task switching in working memory: Effects of performance and aging. *Psychology, 98,* 2095-2100.

Snell, K. B. (1997). Age-related changes in temporal gap detection. *Journal of the Acoustical Society of America, 101,* 2214-2220.

Spilich, G. J. (1983). Life-span components of text processing: Structural and procedural differences. *Journal of Verbal Learning and Verbal Behavior, 22,* 231-244.

Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking and speech understanding in noise in the elderly. *Journal of speech and hearing research, 35,* 1410-1421.

Trainor, L. J., & Trehub, S. E. (1989). Aging and auditory temporal sequencing: Ordering the elements of repeating tone patterns. *Perception & Psychophysics, 45,* 417-426.

Tremblay, K. L., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology, 114,* 1332-1343.

Tun, P. A., & Wingfield, A. (1999). One voice too many: adult age differences in language processing with different types of distracting sounds. *Journal of Gerontology: Psychological Sciences, 54,* 317-327.

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in younger and older listeners. *Psychology and Aging, 17,* 453-467.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary. *Psychological Review, 114*(1), 104-32.

Vanneste, S., & Pouthas, V. (1999). Timing in aging: The role of attention. *Experimental Aging Research, 25,* 49-67.

Verhaeghen, P., & Salthouse, T.A. (1997). Meta-Analyses of age cognition relations in Adulthood: Estimates of Linear and Nonlinear Age Effects and Structural Models. *Psychological Bulletin, 122,* 231-241.

Wong, P. C., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia, 47,* 693-703.

Wright, R. (1981). Aging, divided attention and processing capacity. *Journal of Gerontology, 36,* 605-614.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In T. A. Salthouse & F. I. M. Craik, F. I. M. (Eds.)*Handbook of Aging and Cognition, 2nd Edition* (pp. 293-357). Mahwah, NJ: Lawrence Erlbaum.

Zec, R. F. (1995). The neuropsychology of aging. *Experimental Gerontology, 30,* 431-442.

Relationship between Speech Evoked ALLR and Dichotic CV Scores in Children with Dyslexia

[1]Rohit Kumar Gupta & [2]Prawin Kumar

Abstract

Children with dyslexia may be associated with (central) auditory processing disorders. These processing disorders of an individual can be assessed either through behavioural tests and/or electrophysiological tests. Behavioural and electrophysiological tests are useful in uncovering the important aspects of neural basis of central auditory dysfunction in children with dyslexia. The present study evaluated the performance of children with dyslexia in comparison to typically developing children on speech evoked ALLR and Dichotic CV test. It was also aimed to see if any correlation exists between speech evoked ALLR and in Dichotic CV test in children with dyslexia. A total number of 30 children in the age range of 10 to 12 years were taken for the study. Out of 30 children, there were 15 typically developing children and 15 children with dyslexia. The results revealed that overall for speech evoked ALLR, latencies were significantly prolonged and amplitude was reduced in dyslexic children as compared to typically developing children. Similarly, Dichotic CV scores were also significantly reduced in dyslexic children. Further, it was observed that there were positive correlation between double corrected score and speech evoked ALLR but statistically non-significant. Hence, it is concluded that children with dyslexia performed poorly in dichotic listening test. It is also concluded that there is abnormal encoding of speech signal at cortical level in these children.

Keywords: *Dyslexia, Speech evoked ALLR, Dichotic listening, Dichotic CV test.*

Introduction

Dyslexia is a specific learning disability that is neurological in origin. It is characterized by difficulties with accurate or fluent word recognition and by poor spelling and decoding abilities. Learning problem is one of the common educational problems seen in a number of school going children. This learning problem negatively affects a variety of behaviours, so early intervention is one of the most important steps in this regard. In India, the occurrence of dyslexia ranges from 3% to 7.5% of children (Ramma, 2000). The prevalence estimate of this disability has been found to be 3 to 10 % (Snowling, 2000). Children with dyslexia may have auditory processing disorder and have been experimentally investigated by many researchers (Bellis, 1996; Billiet & Bellis, 2011; Johnson, Nicol & Kraus, 2005; Kraus et al., 1996; Rosen & Manganari, 2001). Studies on incidence of auditory processing deficits in children with dyslexics are estimated to be of 40% (Ramus, 2003).

Studies have shown abnormal processing of speech stimuli and normal processing for tonal stimuli in dyslexic children (Serniclaes, Sprenger-Charolles, Carre & Demonet, 2001). Tallal (1980) reported that there is deficit in processing of brief, rapidly changing auditory stimulus in dyslexics. Study has suggested that such children have difficulty in processing of complex stimuli especially to process through auditory mode (Estes & Huizinga, 1974; Manson & Mellor, 1984).

[1]Email: rht.aslp@gmail.com,
[2]Lecturer in Audiology, Email: prawin_audio@rediffmail.com

The auditory processing of an individual can be assessed through behavioural tests or either by electrophysiological tests. Behavioural and electrophysiological tests have also been useful in uncovering the important aspects of neural basis of central auditory dysfunction. Behavioural tests mainly cut down the external redundancy and assess processing of auditory signal. These behavioural tests includes Dichotic tests, Competing sentence test, Staggered spondaic word test, Pitch pattern test, Duration pattern test, and Gap detection test. These tests are clinically useful to assess one or more auditory processes like auditory integration, sequencing, attention etc.

Dichotic listening test has been frequently used to evaluate the binaural separation and binaural integration abilities. Dichotic CV test has been used to evaluate the normal and impaired auditory process at the cortical level. Dichotic speech test includes variety of stimuli such as nonsense syllables, digits, monosyllabic words, spondaic words or sentences. Studies have shown that children with dyslexia exhibit poorer dichotic listening abilities (Moncrieff & Musiek, 2002; Purdy, Kelly & Davies, 2002).

On the other hand, electrophysiological tests assess the underlying physiology of the auditory system. Auditory evoked potentials provide strong objective methods to assess the neural integrity of the auditory pathway from auditory nerve to cortex (Hood, 1998). Majority of electrophysiological tests has been carried out in individuals with learning disability to assess the auditory processing at the cortical level. The Auditory long latency response (ALLR) is the most frequently used test among the cortical potentials to assess cortical region. Most of the studies have reported a prolonged latency (Arehole, 1995; Guruprasad, 1999; Jirsa & Clontz, 1990; Radhika, 1997) and reduced amplitude in these populations (Jirsa & Clontz, 1990; Mason & Mellor, 1984; Radhika, 1998). David and Ghosh (1984) recorded P1, N1, P2 and N2 peaks in individuals with reading problem and results reveal an increased latency of P1 and P2 peaks when compared with normal average readers. Arehole (1995) studied the relationship between long latency responses and learning disorders in individuals with dyslexia. Results revealed an increased P2-P1 inter-peak latency in individuals with dyslexia in comparison to normal children.

Johnson et al. (2005) described that the synthetic /da/ syllable has been used to study the processing of complex stimuli like speech, at the level of brainstem as well as at the level of cortex and further to study deviancies if any, in clinical population like learning disability. The response manifests as a series of brief neural events that are time-locked to the onset, offset, and the sustained information of the stimulus /da/. This tool has been used to assess binaural listening processing in children with learning disability including dyslexia. Therefore it has been suggested that the use of speech evoked ALLR in assessing such kind of processing deficits is promising to be a valid and reliable tool in such clinical population.

Though behavioural tests have been widely accepted to be the test of choice, however processing deficits may be co-morbid with a number of the pathologies that prevent the administration of behavioural tests. Hence, an attempt is required to check the equivalency of electrophysiological tests in the assessment of (central) auditory processing disorders in children with dyslexia. Cortical potentials have been widely used to understand the neurophysiological basis for speech perception, which would give information of speech processing abilities of the individuals. One such potential may be speech evoked ALLR.

Speech evoked ALLR helps in assessing the capacity of auditory cortex to detect changes within the speech stimuli (Martin & Boothroyd, 1999). There are different types of speech signals which are quite useful in eliciting ALLR includes natural or synthetic vowels, syllables and words (Ceponieni et al., 2001; Sharma, Marsh & Dorman, 2000; Tremblay Friesen, Martin & Wright, 2003). Hence, the recording of ALLR using speech stimuli can probe how the brain processes the signals that underlie auditory detection and discrimination. Majority of the studies have focused on recording of ALLR on click stimulus or more frequency specific tone bursts. But recording of ALLR using tone burst does not give much information about the processing or perception of speech. The P1-N1-P2 evoked neural response is heavily influenced by acoustic content of evoking signal. Hence it is important to know more about how the speech signal is processed in children with dyslexia. Therefore, the speech stimuli /da/ was used in the present study.

Most of the studies have evaluated children with dyslexia, and they observed clinically significant reductions in dichotic listening performance (Maerlender, Wallis & Peter, 2004; Moncrieff & Black, 2008). However, research done in dichotic listening test is very limited in clinical population such as dyslexia, where it can be used as a tool to identify the individuals with dyslexia. Hence, the aim of the present study was to evaluate the efficacy of speech evoked ALLR and dichotic CV test in individuals with dyslexia in comparison to typically developing children. It was also aimed to find out if any correlation exists between speech evoked ALLR and Dichotic CV tests in children with dyslexia.

Method

Participants

Two groups of participants were included in the study: control and experimental group. Thirty participants (60 ears) from both the groups in the age range of 10 to 12 years participated in the study. Control and experimental group consisted of 30 ears from 15 typically developing children (13 males & 3 females) and 15 children with dyslexia (11 males & 4 females) respectively. The diagnosis for the experimental group was made by speech language pathologists / Psychologists at AIISH, Mysore. All the participants had hearing sensitivity within normal limits (hearing threshold less than 15 dB HL at octave intervals between 250 Hz to 8000 Hz for air conduction and between 250 Hz to 4000 Hz for bone conduction), normal middle ear functions as per immittance evaluation, and average or above average intelligence, based on Raven's progressive matrices were selected for the study. However, those participants who were diagnosed as dyslexia with any additional associated problems such as attention deficit disorder with/without hyperactivity, chronic psychological disorder, or with any other neurological disorder were excluded from the study.

Instrumentation

A calibrated two channel diagnostic audiometer (Orbiter-922) with TDH-39 headphones and MX-14/AR ear cushion was used for air conduction thresholds. Radio ear B-71 bone vibrator was used for estimating bone conduction thresholds. The same audiometer was used for presenting dichotic CV test stimuli coupled with personal computer. A calibrated middle ear analyzer (GSI-Tympstar, version 2) was used to rule out middle ear pathology. ILO version 6 was used to record the TEOAEs. Bio-logic Navigator pro (ver-

Table 1: Protocol for recording click evoked ABR and speech evoked ALLR

Parameters	Click evoked ABR	Speech evoked ALLR
Stimulus	Click (100 μs duration)	Natural /da/ stimulus (185 ms)
Electrode Placement	Non-inverting-Fpz Common-A1/A2 Inverting-A2/A1	Non-inverting-Fpz Common-A1/A2 Inverting-A2/A1
Intensity	90 dBnHL	80 dBnHL
Polarity	Rarefaction	Alternating
Filter setting	100 - 3000 Hz.	1 - 30 Hz.
Repetition rate	30.1/sec	1.1/sec
Time window	10-12 ms	500 ms
No. of channel	Single	Single
No. of sweeps	1500	200
Impedance	$< 5\text{k}\Omega$	$< 5\text{k}\Omega$
No. of replication	2	2

sion 7.0) evoked potential system was used for recording click evoked auditory brainstem response (ABR) and speech evoked ALLR.

Test Materials

For speech evoked ALLR, a natural /da/ stimulus was recorded by an adult male speaker with clear articulation. The recording was done using unidirectional microphone connected to the computer in the sound treated room. Adobe Audition (version 2) software with a sampling rate of 48000 Hz and 16 bit resolution was used. The stimulus duration was approximately 185 ms. Recorded stimulus was then converted into wave file and loaded into the Biologic navigator pro evoked potential system for speech evoked ALLR recording.

For dichotic CV test, the material used was dichotic Consonant-Vowel (CV) word lists (Yathiraj, 1999) consisting of 30 standardized pairs of syllables /pa/, /ta/, /ka/, /ba/, /da/ and /ga/.

Test Environment

The testing was carried out in an acoustically sound treated room with ambient noise levels within permissible limits as per ANSI S3.1 (1991).

Test Procedure

Screening Checklist for Auditory Processing (SCAP) was administered on control group developed by Yathiraj and Mascarenhas (2003), to rule out symptoms of auditory processing disorders. It consists of twelve questions having the symptoms of deficits in auditory processing. The scoring was done on a two point rating scale (Yes/No). Children who scored less than 50% were considered for the study.

Pure tone thresholds were obtained at octave intervals between 250 Hz and 8000 Hz for air conduction and between 250 Hz and 4000 Hz for bone conduction (mastoid placement), using modified Hughson and Westlake procedure (Carhart & Jerger, 1959). Tympanometry was carried out using 226 Hz probe tone at 85 dBSPL to rule out any middle ear pathology. For reflexometry, acoustic reflex measurement was performed using reflex eliciting tone of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz ipsilaterally and contralaterally. Transient evoked otoacoustic emissions (TEOAE) were measured using click stimuli at 85 dBSPL in both ear to assess the outer hair cells functioning.

Click evoked ABR and speech evoked ALLR were recorded in both ears for all the participants using the test protocol mentioned in Table 1. Participants were made to sit comfortably in order to ensure a relax posture and minimum rejection rate. Gold cup electrodes were placed after cleaning the electrode placement sites with preparing gel. Conduction paste was used to improve the conductivity of the recording signal from the generator sites. The electrodes were secured to the place by using plasters. The electrode placement was kept and followed as per the test protocol.

Dichotic CV test: The dichotic consonant-vowel test material was played through personal computer connected to the calibrated double channel diagnostic audiometer. The dichotic CV word lists were presented to both the ears using zero (0) ms lag at 40 dB SL (re: SRT). The children were instructed as "*You will be hearing two words one to each ear at the same time. You should repeat both the words that you hear*". Task understanding was ensured using five practice items before proceeding to the dichotic consonant-vowel test.

Table 2: Mean and standard deviation of latency and amplitude measure for control and experimental groups

Waves	Latency Measures (ms)				Amplitude measures(μV)			
	Control group (N = 29)		Experimental group (N = 26)		Control group (N = 29)		Experimental group (N = 26)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
P1	91.26	15.52	113.35	31.57	2.88	1.23	1.41	0.82
N1	134.48	23.67	175.14	37.15	-2.15	1.51	-3.88	1.96
P2	192.43	42.57	242.65	43.95	1.99	0.81	0.71	0.52
N2	235.68	46.09	303.20	39.19	-1.48	1.41	-2.81	1.40

N = number of ears; SD= Standard deviation; ms = millisecond; μV = microvolt.

The responses of all the participants were scored in terms of single correct score and double correct scores. The right ear score (RES), left ear score (LES), and double correct score (DCS) were scored. A single correct score was calculated when the participants reported the syllable presented to any one ear correctly. A double correct response was calculated when the participants reported the syllable presented to both ears correctly.

Statistical Analysis

Descriptive statistical analysis of the scores in terms of mean, standard deviation and other tests (parametric and non-parametric) such as Multivariate analysis of variance (MANOVA), Paired t-test and Karl Pearson correlation test performed using Statistical package Social Science (SPSS 16.0) software for both speech evoked ALLR and Dichotic CV tests. The results obtained are presented and discussed in the subsequent section.

Results and Discussion

Descriptive statistics was done to find out the mean and standard deviation (SD) for all the parameters for both control and experimental groups. MANOVA was administered to compare between experimental as well as control group for latency and amplitude of speech evoked ALLR and behavioural tests (dichotic CV) scores. In order to find out ear advantage Paired t-test was used to compare between the two groups. Karl Pearson's Correlation was done to check whether any relationship exists between speech evoked ALLR and DCS of dichotic CV scores.

Speech Evoked ALLR

In typically developing children all the peaks of speech evoked ALLR was present in all participants. However, in children with dyslexia, all peaks of speech evoked ALLR was present in all participants expect wave N2, which was absent in two participants. Hence, the percentage of recorded waveform for P1, N1, and P2 was 100% whereas for N2 it was only 86.6%. The mean and standard deviation of speech evoked ALLR latency and amplitude measures are mentioned in Table 2.

*P1 wave:*From Table 2, it can be observed that the mean latency of P1 was significantly prolonged and amplitude was reduced in dyslexic children as compared to typically developing children ($p < 0.05$). This is in agreement with the findings of Satterfield et al. (1984) and Byring and Jaryilehto (1985). They also reported that P1 latency was delayed and amplitude was reduced in children with learning disability. They attributed it could be probably due to delayed maturation in children with learning disability.

N1 wave: From the visual inspection it was observed that wave N1 present in all dyslexic children. This finding was consistent with the findings obtained by Radhika (1998). They found that N1 was present in all children with learning disability. Moreover, the results of MANOVA (Table 3) revealed that latency of N1 was significantly prolonged in children with dyslexia as compared to typically developing children. ($p < 0.05$). The amplitude of N1 was also found to be significantly reduced in dyslexic children as compared to typically developing children. These findings are consistent with the finding of other's researchers (David & Ghosh., 1984; Kibble et al., 1986; Pinkerton et al., 1989). They reported that the latency was increased and amplitude was reduced in children with learning disability. This could be related to short attention span in children with dyslexia (Picton et al., 1978). This suggests that children with learning problem take longer time to initiate the negativity.

Wave P2: As reflected from Table 2, latency of wave P2 was significantly prolonged and amplitude was reduced in children with dyslexia as compared to typically developing children. ($p < 0.05$). The similar results of increased latency and reduced amplitude of waves P2 was reported by David and Ghosh (1984), Byring and Jaryilehto (1985). They found that the latency of wave P2 was prolonged and amplitude was reduced in children with learning disability as compared to children without learning problem.

Wave N2: From the visual inspection it was observed that out of 15 dyslexic children wave N2 was absent in two participants. In rest of the children with dyslexia the wave N2 was quite visible. This finding was con-

Table 3: F-value for latency and amplitude measure between control and experimental groups

Waves	Latency measures		Amplitude measures	
	F-value	p-value	F-value	p-value
P1	F (1,53) = 11.42	0.001**	F (1,53) = 13.04	0.001***
N1	F (1,53) = 23.92	0.000***	F (1,53) = 6.59	0.017*
P2	F (1,53) = 18.50	0.000***	F ((1,53) =23.19	0.000***
N2	F (1,53) = 33.83	0.000***	F ((1,53) = 5.99	0.022*

*p<0.05; **p<0.01; ***p<0.001

sistent with the findings obtained by Radhika (1998). They found that N2 was absent in 8 children out of 12 children with learning disability. Moreover, the results of MANOVA (Table 3) reveled that latency of N2 was significantly prolonged in children with dyslexia as compared to typically developing children. ($p < 0.05$). There are few studies have been reported in literature on latency of N2 because of wide range of latency variation observed in normal individuals. In the present study these deviancy in N2 latency could be because of the deficits in auditory processing of temporal aspects of the stimuli, which required more controlled attention (David et al., 1984; Byring & Jaryilehto, 1985).

To conclude, present study finding suggests that children with dyslexia performed poorly in comparison to typically developing children. This outcome is based on the differences observed in latency and amplitude measures between two groups in speech evoked ALLR.

Dichotic Consonant-Vowel (CV) Test

Descriptive statistics was done to obtained mean and standard deviation (SD) of dichotic CV score in terms of single correct scores (SCS) of right and left ear and double correct scores (DCS) for children with dyslexia and a typically developing children. *MANOVA* results showed thatthere were statistically differences between two groups on behavioural dichotic CV scores (Table 4 & 5). Mean scores of typically developing children were much poorer in comparison to dyslexic children. This difference in performance for left ear and right ear single correct score between the two groups were statistically significantly ($p < 0.05$). Similarly double corrected scores (DCS) between the two groups was also statistically significant ($p < 0.05$).

These findings suggest that children with dyslexia shows binaural integration deficit in comparison to typically developing children. The present finding is in agreement with other studies on dichotic listening test which assessed binaural integration processes (Ayers, 1972; Billiet & Bellis, 2011; Cermak & Koomar, 1981; Helland, Asbjornsen, Hushovd & Hugdahl, 2008; Moncrieff & Black, 2008; Mortan & Siegel, 1991). These studies reported that children with learning disability scored significantly poorer in comparison with typically developing children. In a similar line, Ganguly, Rajagopal and Yathiraj (1994) also found reduced scores in children with learning disability on the dichotic CV test in comparison to typically developing children.

Moreover in order to find out the ear advantage the *paired t-test* was used to compare right and left ear of control as well as experimental group. These differences between the ear scores were statistically significant ($p < 0.05$), which showed right ear advantage in typically developing children.

On the other hand, experimental group shows significantly higher scores for left ear ($p < 0.05$), which pointed towards left ear advantage in children with dyslexia. These finding in children with dyslexia in comparison to typically developing children revealed the differences in performance between two ears (ear advantage) and heterogeneity among dyslexic groups. These discrepancies probably are because of differences in degree of reading and writing impairment (Helland et al., 2008; Hugdahl et al., 1995; Iliadou, Kaprinis, Kandylis & Kaprinis, 2010; Moncrieff & Black, 2008). The lack of right ear advantage (REA) in children with dyslexia observed in present study may also be explained by a tendency seen in dyslexics to switch attention between the ears rather than splitting their attention (Dickstein & Tallal, 1987). In the another study Iliadou et al. (2010) found that the purely dyslexic children shows an almost equally distributed right and left hemispheric dominance as well as no dominance of the two hemispheres. This shows heterogeneity of dyslexic group in comparison to typically developing children.

To conclude, behavioural dichotic CV tests results in present study revealed poorer performance for children with dyslexia in comparison to typically developing children. The poorer performance was observed for single correct scores as well as for double correct scores between two groups. It was also observed that typically developed children show REA whereas children with dyslexia exhibit left ear advantage (LEA). Hence, the outcome of present study indicates there is binaural integration deficit in children with dyslexia in comparison to typically developing children.

Table 4: Mean and standard deviation values of Dichotic CV scores in control and experimental group

Scores	Control group (N = 15)		Experimental group (N = 15)	
	Mean*	SD	Mean*	SD
LCS	16.93	1.71	14.07	2.31
RCS	22.80	1.56	11.93	2.25
DCS	10.93	2.37	4.93	1.38

**Maximum scores=30; LCS=left correct scores; RCS=right correct scores; DCS=double correct scores; N = number of participants*

Table 5: F values for Dichotic CV scores between control and experimental group

Scores	F-value	p- value
LCS	F (1,28) = 14.89	0.001***
RCS	F (1,28) = 235.42	0.000***
DCS	F (1,28) = 71.41	0.000***

p<0.05; **p<0.01;*p<0.001*

Relationship between Speech-evoked ALLR and Dichotic CV Test

Karl Pearson's correlation was used to check whether there were any relationship between different components of speech evoked ALLR (latency & amplitude) and double corrected scores of dichotic CV tests in typically developing children as well as dyslexic children.

From the Table 6 it can be inferred that there is a positive correlation between dichotic listening and speech evoked ALLR. This suggest that dichotic listening tends to be poorer when there is an increase (prolong) in latency and reduction in amplitude of ALLR and vice versa. Though, there was the relationship between speech evoked ALLR and dichotic listening, it was not statistically significant. This suggests that the speech evoked ALLR could be taken as alternative tool to assess dichotic listening. Moreover, we did not find any study in this regard to support the findings of present study. Probably larger sample size may be required to validate the present findings.

Table 6: Correlation between Speech evoked ALLR and Dichotic CV test

Parameters	Double correct scores			
	Latency Measures		Amplitude measures	
	correlation coefficient (r)	p-value	correlation coefficient (r)	p-value
P1	0.23	0.39	0.41	0.60
N1	0.06	0.80	0.10	0.71
P2	0.00	0.99	0.02	0.92
N2	0.32	0.28	0.47	0.10

However, there are some studies which indirectly support the present findings. Banai et al. (2009) studied the relationship between sub-cortical auditory encoding and literacy-related skills in children with learning problems. They observed statistically significant correlation between the measures of timing components (transition) of sub-cortical auditory encoding and reading skills. Moreover, the relationship was less significant between the harmonic components (formants) of sub-cortical auditory encoding and reading skills. Based on these findings they concluded that good readers show more temporally precise encoding and more robust representation of speech harmonics in comparison to poor readers who represent poor timing of sub-cortical auditory encoding and impoverished representation of signal harmonics. To conclude, there is a relationship between dichotic listening and speech evoked ALLR. However it was not statistically significant in children with dyslexia.

Conclusions

Speech evoked ALLR is easily traceable in all children with dyslexia as well as typically developing children. The results of speech evoked ALLR indicates abnormal encoding of speech signal at the cortical level in children with dyslexia. Further, Dichotic CV test showed poorer dichotic listening ability for dyslexic children in comparison to typically developing children. In addition, correlation analysis suggest that dichotic listening tends to be poorer when there is increase in latencies or reduction in amplitude of speech evoked ALLR in dyslexic children. Hence, from the above findings present study highlights the importance of speech evoked ALLR as tool to assess dichotic listening along with dichotic test.

References

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

Arehole, S., (1995). A Preliminary Study of Relationship between Long Latency Response and Learning Disorder. *British Journal of Audiol-*

ogy, 29, 295-298.

Ayers, A. (1972). *Sensory Integration and Learning Disorders.* Los Angeles: Western Psychological Services.

Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex, 19*, 2699-2707.

Bellis, T. J. (1996). *Assessment and Management of Central Auditory Processing Disorders in Educational Settings: From Science to Practice*: Singular Publishing Group, Inc: San Diego.

Billiet, C. R., & Bellis, T. J. (2011). The Relationship between Brainstem Temporal Processing and Performance on Tests of Central Auditory Function in Children with Reading Disorders. *Journal of Speech, Language, and Hearing Research, 54,* 228-242.

Byring, R., & Jaryilehto, T. (1985). Auditory and Visual Evoked Potential of Schoolboys with Spelling Disabilities. *Developmental Medicine and Child Neurology, 27,* 141-148.

Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders, 24*, 350-345.

Ceponiene, R., Shestakova, A., Balan, P., Alku, P., Yiaguchi, K., & Näätänen, R. (2001). Children's auditory event-related potentials index sound complexity and "speechness". *The International Journal of Neuroscience, 109,* 245-260.

Cermak, S. A., & Koomar, J. A. (1981). Reliability of dichotic listening using two stimulus formats with normal and learning-disabled children. *The American Journal of Occupational Therapy, 35,*456-463.

David, J., & Ghosh, M. (1984). Late Latency Response in Poor Readers. *Electroencephalography and Clinical Neurophysiology, 70,* 232-237.

Dickstein, P., & Tallal, P. (1987). Attentional capabilities of reading impaired children during dichotic presentation of phonetic and complex nonphonetic sounds. *Cortex, 23*, 237-249.

Estes, R. E., & Huizinga, R. J. (1974). A Comparison of Visual and Auditory Presentation of a Paired-Associate Learning Tests with Learning Disabled Children. *Journal of Learning Disabilities, 7,* 44-51.

Ganguly, L., Rajagopal, L., & Yathiraj, A. (1994). Perception of Dichotic CV Syllables in Normal and Learning Disabled Children. *Journal of Indian Speech and Hearing Association, 13,* 56-59.

Guruprasad, A. (1999). Evaluation of central auditory processing disorder in children with learning disability. *Unpublished Masters Dissertation,* University of Mysore, Mysore.

Helland, T., Asbjornsen, A. E., Hushovd, A. E., & Hugdahl, K. (2008). Dichotic Listening and School Performance in Dyslexia. *Dyslexia, 14*, 42-53.

Hood, L. J. (1998). *Clinical applications of auditory brainstem responses.* Singular publishing group, San Diego Inc.

Hugdahl, K., Helland, T., Farevaag, M. K., Lyssand, E. T., & Asbjornsen, A. (1995). Absence of ear advantage on the consonant-vowel dichotic listening test in adolescent and adult dyslexics: Specific auditory-phonetic dysfunction. *Journal of Clinical and Experimental Neuropsychology, 17,* 833-840.

Iliadou, V., Kaprinis, S., Kandylis, D., & Kaprinis, J. S. (2010). Hemispheric laterality assessment with dichotic digits testing in dyslexia and auditory processing disorder. *International Journal of Audiology, 49*, 247-252.

Jirsa, R. E. & Clontz, K. B. (1990). Long Latency Auditory Event Related Potentials from Children with Auditory Processing Disorders. *Ear and Hearing, 11,* 222-232.

Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brainstem response to speech: A biological marker of auditory processing. *Ear and Hearing, 26,* 424-434.

Katz, J., Burkard, R. F., & Medwetsky, L. (2002). *Handbook of Clinical Audiology* (5^{th} Ed.), Lippincott Willliams & Wilkins.

Kibble, K., Michael, M. A., Suzanne, B., Verkert, M. A., Karen, M. G., & Frank, E. M. (1986). Late Latency Potentials and P300. *Hearing Instruments, 37,* 22-24.

Kraus, N., McGee, T., Carell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996). Auditory Nerophysilogical Responses and Discrimination Deficits in Children with Learning Problems. *Science, 273,* 971-973.

Maerlender, C., Wallis,D. J. & Peter, K. (2004). Psychometric and Behavioral Measures of Central Auditory Function: The Relationship between Dichotic Listening and Digit Span tasks. *Child Neuropsychology, 10*, 318-327.

Manson, S. M., & Mellor, D. H. (1984). Brainstem, middle latency and late cortical evoked potentials in children with speech and language disorders. *Electroencephalography and Clinical Neurophysiology, 59,* 297-309.

Martin, B. A., & Boothroyd, A. (1999). Cortical, auditory, event related potentials in response to periodic and aperiodic stimuli with the same spectral envelop. *Ear and Hearing, 20,* 33-44.

Moncrieff, D. W., & Black, J. R. (2008). Dichotic Listening Deficits in Children with Dyslexia.

Dyslexia, 14, 54-75.

Moncrieff, D. W., & Musiek, F. E. (2002). Interaural asymmetries revealed by dichotic listening tests in normal and dyslexic children. *Journal of the American Academy of Audiology, 13,* 428-437.

Morton, L. L., & Siegel, L. S. (1991). Left ear dichotic listening performance on consonant vowel combinations and digits in subtypes of reading-disabled children. *Brain and Language, 40,* 162-180.

Picton, T. W., Woods, D. L., & Proulx, G. B., (1978). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology, 45,* 198-210.

Pinkerton, F., Watson, D. R. & McClelland, R. J. (1989). A Neurophysiological Study of Children with Reading, Writing and Spelling Difficulties. *Developmental Medicine and Child Neurology, 31,* 569-581.

Purdy, S. C., Kelly, A. S., & Davies, M. G. (2002). Auditory brainstem response, middle latency response, and late cortical evoked potentials in children with learning disabilities. *Journal of the American academy of audiology, 13,* 367-82.

Radhika, A. (1998). Auditory Late Latency Potentials in Learning Disabled Children. *An Unpublished Independent Project, University of Mysore,* Mysore.

Ramma, S. (2000). Two decades of research on learning disabilities in India. *Dyslexia, 6,* 268-283.

Ramus, F. (2003). Developmental Dyslexia: Specific phonologic deficit or general sensorimotor dysfunction? *Current opinion in Neurobiology, 13,* 212-218.

Rosen, S., & Manganari, E. (2001). Is There a Relationship Between Speech and Nonspeech Auditory Processing in Children with Dyslexia? *Journal of Speech, Language and Hearing Research, 44*, 720-736.

Satterfield, J. H., Schell, A. M., Backs, R. W., & Hidaka, K. C. (1984). Cross sectional and Longitudinal Study of Age Effects of Electrophysiological Measures in Hyperactive and Normal Children. *Biological Psychiatry, 19,* 973-990.

Serniclaes, W., Sprenger-Charolles L., Carre R., & Demonet, J. F. (2001). Perceptual Discrimination of Speech Sounds in Developmental Dyslexia. *Journal of Speech, Language and Hearing Research, 44,* 384-389.

Sharma, A., Marsh, C. M., & Dorman, M. D. (2000). Relationship between N1 evoked potential morphology and the perception of voicing. *Journal of the Acoustical Society of America, 108,* 3030-3035.

Snowling, M. J. (2000). *Dyslexia (2^{nd} Ed.),* Oxford: Blackwell Publishers Ltd.

Tallal, P. (1980). Auditory temporal perception, phonics and reading children. *Brain and Language, 9,* 182-198.

Tremblay, K. L., Friesen, L., Martin, B. A., & Wright, R. (2003). Test retest reliability of cortical evoked potentials using naturally produced speech sounds. *Ear and Hearing, 24,* 225-232.

Yathiraj, A. (1999). *Dichotic CV test- revised.* Developed at Department of Audiology, All India Institute of Speech and Hearing, Mysore.

Yathiraj, A., & Mascarenhas, K. E. (2003). Effect of auditory stimulation in central auditory processing disorder. *AIISH Research Fund Project*, All India Institute of Speech and Hearing, Mysore.

The Effect of Digital Noise Reduction (DNR) in Hearing Aids on Auditory Late Latency Response, Speech Recognition Ability and Quality

[1]Sahana, P. & [2]Manjula, P.

Abstract

The present study aimed to investigate the effect of Digital noise reduction (DNR) on auditory late latency response (ALLR), speech recognition ability (SRS) and quality of speech. The data were collected from 10 individuals with normal hearing and 14 individuals with hearing loss. The results revealed that there was a negative effect of noise on the SRS, latencies of ALLR components (P1, N1 & P2) and the amplitude of N1-P2 complex. The results also suggest that there was a slight improvement in SRS on activation of DNR. The improvement brought about by the DNR was not enough to bring the scores to that obtained under quiet condition. The DNR activation minimizes the effect of background noise by reducing the prolongation of the latencies of ALLR peaks. Nonetheless, the amplitude of N1-P2 complex remained unchanged by the activation of DNR. Further, there was no relationship between the SRS and N1-P2 amplitude in quiet and noise. However, there was no correlation between the morphology ratings and SRS obtained under DNR activated and deactivated condition. In addition, DNR significantly improved the 'Loudness','Clarity', 'Naturalness' and 'Overall impression' ratings for the speech through the hearing aid. The quality ratings also appears to be 'acceptable to excellent reliability' between the two sessions on the three parameters namely 'Loudness', 'Clarity' and 'Overall impression'. Overall, the effect of white noise was greater for individuals with hearing loss. Participants with moderate hearing loss are affected by noise to a greater degree compared to those with mild hearing loss.

Keywords: *DNR, ALLR, SRS and quality of speech.*

Introduction

Hearing loss and other perceptual problems related to aging cause communicative difficulties (Gelfand, Piper, & Silman, 1986; Nabelek, 1988). Due to this communication difficulty, reduced psychosocial function has often been reported. In particular, there is a decline in social interaction, intimate relations, self-concept, psychological status, and cognition (Weinstein & Ventry, 1982; Scherer & Frisina, 1998). Majority of persons with mild-to-moderate hearing loss indicate that their primary problem is difficulty hearing in noise (Kochkin, 2005). Thus, listening in background noise presents a challenge that often leads to communication breakdowns.

In addition, successful communication in difficult listening environments will depend on how the auditory system is able to extract signals of interest from other competing information. Thus, cortical auditory evoked potential (CAEP) is another approach to study the encoding of signal in noise, in the human central auditory system (CAS). It is a measure of CAS function that can provide valuable information about the way in which neurons encode signals in noise (Billings, Tremblay, Steckera, & Tolina, 2009).

There are a few studies which have recorded the CAEPs in individuals with normal hearing. However, there is a dearth of literature on the cortical auditory evoked potentials associated with encoding signal in individuals with hearing loss. The results of such studies indicate that the latencies are more sensitive indicators of these masking effects than amplitudes (Whiting, Martin, & Stapells, 1998; Billings et al, 2009). Morphology of the P1-N1-P2 complex was driven primarily by SNR, highlighting the importance of noise when recording CAEPs. Since SNR is the one that determines the efficiency of DNR, CAEPs can also be used as a measure to evaluate this aspect.

Further, it is reported that the components of auditory evoked late latency response (ALLR) can be correlated with the behavioral measures of speech perception in quiet (Narne & Vanaja, 2008). Chandra and Barman (2009) investigated the relationship between the late latency response and the speech identification scores, in noise at 0 dB SNR, for different speech stimuli (/da/, /ba/ and /ga/) in persons with auditory neuropathy. The results revealed that there was no correlation between the amplitude and latency of the potentials and the speech identification scores (SIS). Hence, there are mixed results seen regarding the correlation between the behavioural measure of speech perception and the ALLR.

The usual remedy for people with cochlear hearing loss is amplification through hearing aids. This hearing aid improves speech perception in quiet conditions mainly by increasing the audibility. However, in the presence of noise, a hearing aid amplifies the background noise as well the speech which causes annoyance due to the amplified background noise. This results in poor speech intelligibility, due to upward spread of masking at high

[1]Email: sahanap77@gmail.com,
[2]Professor in Audiology, Email: manjulap21@aiishmysore.in

listening levels, distortion caused by limited bandwidth of the hearing aid (Plomp, 1978). Therefore, there are a variety of signal-processing techniques in hearing aids to tackle this problem. The commercially available hearing aids have different algorithms to improve signal-to-noise ratio such as Digital Noise Reduction (DNR) and directional microphone. The goal of these is to improve speech intelligibility in noise or to provide comfort in noisy situations or both.

DNR in hearing aid has a general goal of providing less amplification over a specified frequency range, for noise than for speech. The DNR algorithm relies on the difference in physical characteristics of a signal to distinguish speech from noise (Ricketts, & Hornsby, 2005). There are studies that have investigated the efficacy of digital noise reduction on the perception of speech embedded in noise, using behavioral measures.

Alcantara, Moore, Kuhnel, and Launer (2003), have evaluated the effectiveness of a noise reduction system implemented in a commercial digital multi-channel compression hearing aid, in individuals with moderate sensori-neural hearing loss. The results reported ratings of sound quality; listening comfort and the SRT were very similar, with and without the noise reduction system. In contrast, Ricketts and Hornsby (2005) reported that their participants showed the strong preference for DNR processing and concluded that implementation of DNR processing improved sound quality but not the speech recognition in speech-in-noise condition. Also the study done by Mueller, Weber, and Hornsby (2006) report that DNR processing will increase ease of listening by reducing the annoyance in speech-in-noise situations. Thus, these studies have shown equivocal results on the sound quality of speech output by the implementation of DNR in hearing aids. These results clearly state that the DNR signal processing will not have an effect on speech understanding in the presence of noise.

However, Bray and Nilsson (2001) reported that for noise arriving from the front condition, the mean aided benefit was 2.6 dB SNR without DNR activated; and 3.5 dB SNR with DNR activated. This led to the conclusion that, DNR algorithms in conjunction with directional microphone may be effective in improving speech perception in noise when the noise field is isotropic. So, the DNR alone will not improve the speech perception but it is useful in conjunction with directional microphone (Nordrum, Erler, Garstecki, & Dhar, 2006).

There are abundant studies done in literature, which have evaluated the change in subjective measures like SRT, SIS and quality of speech output on DNR signal processing (Boymans & Dreschler, 2000; Alcantara et al, 2003). However, there is a dearth of literature that reports on the effect of DNR in hearing aids on electrophysiological measures.

The studies revealed that the cortical potentials are sensitive to the relative intensity of the signal with respect to noise rather than the absolute level (Kaplan-Neeman, Kishon-Rabin, Henkin & Muchnik 2006; Whiting et al, 1998; Billings et al, 2009; Chandra & Barman, 2009). A few studies have reported that cortical responses may differ from expected when recorded via sensory devices (e.g., hearing aids), this has given an impetus for additional studies examining auditory evoked responses recorded with hearing aids.

However, there is a dearth of studies in background noise done in population with perceptual difficulties such as those with hearing impairment or older adults. CAEPs may give potential information on the signal-in-noise difficulties experienced by these groups. The findings of such a study would improve one's understanding on how the human auditory cortex encodes signal in noise, in individuals with hearing impairment. Thus, the current study was taken up to explore whether the signal-to-noise ratio improved by the DNR is measurable at the level of cortex.

The primary purpose of the study was to compare the aided performance in terms of speech recognition scores (SRS), auditory late latency response (ALLR) and the quality of speech of the participants, with and without the digital noise reduction (DNR) being activated. Another purpose of the study was to compare the effect of masking noise in individuals with normal hearing and hearing loss on SRS and components of ALLR.

Method

The null hypothesis of the present study was that there was no significant difference between the SRS, ALLR and quality of speech output with DNR activated and deactivated. Repeated measures research design was used to test the null hypothesis.

Participants

The data were collected from a total of 24 participants. All the participants were native speakers of Kannada language (Dravidian language spoken in southern part of India). The participants did not have any psychological and neurological problems. They did not have middle ear pathology as confirmed by immittance evaluation. The participants were divided into two groups; Group A and Group B.

Group A: A total of 10 participants (N=10) were included in the group. The age of the participants ranged from 19 to 40 years (mean age of 27.90 years). The participants in this group had pure tone thresholds within 15 dB HL at octave frequencies between 250 Hz and 8 k Hz. They had $\geq$ 80% speech recognition scores (SRS) in quiet and $>$ 60% speech recognition scores at 0 dB SNR on phonemically balanced bi-syllabic word list in

Kannada (Yathiraj & Vijayalaksmi, 2005).

Group B: The participants in Group B had acquired hearing loss with adequate speech and language. The participants had flat sensorineural hearing loss (SNHL), with air-bone gap not greater than 10 dB. The difference between the highest and the lowest air-conduction threshold across frequency from 250 Hz to 8000 Hz did not vary more than 20 dB from each other (Pittman & Stelmachowicz, 2003). Their SRS was proportionate to the hearing loss (Vanaja & Jayaram, 2005). The Group B participants were further distributed into two; Group B1 and Group B2, based on the degree of hearing loss. Group B1 included 7 participants with mild flat sensorineural hearing loss in the age range of 35 to 55 years (mean age of 44.86 years), and Group B2 also comprised of 7 participants with moderate flat sensorineural hearing loss. Their age ranged from 30 to 55 years (mean age of 42.4 years).

Instrumentation

A calibrated two-channel diagnostic audiometer Madsen OB922 (version 2) with TDH-39 headphones housed in MX-41/AR ear cushions and a bone vibrator, Radio ear B-71 was used to carry out pure tone audiometry. A loudspeaker (Martin Audio, C115) placed at 45 degree azimuth, at one meter distance from the aided ear of the participant, was used for presenting the test stimuli. A calibrated GSI-Tympstar (version 2) immittance meter was used to rule out middle ear pathology. Bio-logic Navigator Pro EP system was used to record ALLRs using dB electronics loudspeaker.

A four channel digital behind-the-ear hearing aid was chosen for the study. According to the manufacturer's specifications, the frequency range of this hearing aid was from 100 Hz to 6800 Hz. A personal computer with NOAH-3 and hearing aid specific software with Hearing instrument Programmer (Hi-Pro) interface were used to program the hearing aids and to activate/deactivate the DNR.

Test material

Phonemically balanced (PB) bi-syllabic word lists in Kannada (Yathiraj & Vijayalaksmi, 2005) were used to find out the speech recognition scores. The judgement of sound quality rating scale was developed by Gabrielsson, Schenkman, and Hagerman (1988), originally with eight dimensions related to sound quality, was adapted for the study.

Recording and Preparation of /da/ Stimulus and Kannada Passage

The Consonant-Vowel (CV) token /da/ was uttered by a female adult speaker, whose mother tongue was Kannada, with normal vocal effort. The /da/ stimulus were recorded in a sound-treated room using the Adobe Audition (Version 1.5) software, installed in the personal computer, via a hand-held unidirectional microphone (AHUJA, AUD-101XLR) placed 10 cm away from the lips of the speaker. The recorded stimulus was digitized using a 32-bit processor at 44,100 Hz sampling frequency. The /da/ stimulus was uttered thrice with an approximate duration of stimulus being 250 ms. Goodness test of /da/ stimulus was carried to see which of the /da/ stimulus was natural, by presenting the stimuli to five individuals with normal hearing. The stimulus with highest rating of goodness was selected.

Likewise, the Kannada passage, picked up by a story was recorded in Abode Audition spoken by an adult female whose mother tongue was Kannada in clear conversational speech style. The passage was given to five individuals with normal hearing for the Goodness test and they rated the passage to be highly intelligible.

All tests were administered in an air-conditioned sound treated double/single room set-up.

Test Procedure

After the audiological evaluation, the participants satisfying the selection criteria were considered for further evaluations conducted in Phases I, II, and III.

Phase I: Fitting and optimizing hearing aid

In this phase, digital behind the ear hearing aid was programmed for each participant in the Group B1 and Group B2, so that the gain was adjusted according to each participant.

The hearing aid was programmed using NOAH and hearing aid specific software on a personal computer. The hearing aid worn by the participant was connected to Hi-Pro through a connecting cable and the hearing aid was detected by the programming software. The hearing thresholds of each participant were fed into the programming software and target gain curves were obtained using the proprietary prescription formula of the hearing aid. The hearing aid gain was first-fit to match the target gain.

After the initial first-fit, the participants were asked to repeat the Ling's six sounds presented randomly (/a/, /i/, /u/, /s/, /sh/ and /m/). The gain was optimized for audibility of the Ling's six sounds by adjusting the gain of the hearing aid until the participants were able to identify all six Ling's sounds. The aided audiogram was also done to ensure adequate audibility.

The hearing aid was set to amplify in omni-directional mode with the volume control deactivated. The hearing aid chosen had two programs. Program 1 of the hearing instrument had speech in quiet program, wherein digital noise reduction was turned 'off'. Program 2 was similar to Program 1 except for the noise reduction algorithm turned 'on'. The settings were saved in the hearing aid for each participant. Finally, the fitting status was saved into the hearing aid. This was repeated for each test ear and for each participant.

Phase II: Behavioural Testing

The Speech Recognition Scores (SRS) and Perceptual quality rating were collected from each test ear of each participant.

Speech Recognition Scores (SRS)

The Speech Recognition Scores (SRS) were obtained using recorded phonemically balanced (PB) word-list in Kannada (Yathiraj & Vijayalakshmi, 2005). The participants were made to sit comfortably on a chair in the test room at a distance of 1 meter and 45^0Azimuth from the loudspeaker of the audiometer. The recorded word list was routed to the loud speaker through the auxiliary input of the audiometer, at 45 dB HL. Before the presentation of the stimuli, the level of the presentation was set to 45 dB HL and level adjustments was done for the calibration tone, such that the VU-meter deflections averaged at 0. The presentation level of the stimuli was monitored with VU meter. The non-test ear was given speech noise of 65 dB HL from the audiometer in order to avoid its participation.

SRS in quiet: The recorded speech material (PB word-list) was presented at 45 dB HL to obtain SRS in quiet, through sound field. The SRS in quiet was obtained for all the participants. For participants in Group A, it was measured in unaided condition; whereas for participants in Group B, the SRS was measured in aided condition. This was measured by presenting one complete PB word-list of 25 words for each condition. The participants were instructed to repeat the words being presented. The responses were scored as the number of words correctly identified. Each correct response was given a score of '1' and each incorrect response was given score of '0'. The maximum score was 25 as each list consisted of 25 words. The total number of correctly repeated words in the list was noted. This was considered as the SRS of the participant for a particular test condition.

SRS in noise: The white noise was calibrated to give same output as speech stimuli, such that routing both speech and noise through the loud speaker would give 0 dB SNR. For obtaining SRS under noise condition, the recorded PB word-list was presented at 45 dB HL and the white noise was also routed through the same loud speaker. The number of words correctly repeated was noted and this gave the SRS under noise condition. SRS in noise was obtained in unaided condition for participants in Group A. For participants in Group B, the SRS were obtained, under two aided test conditions, i.e., by activating and deactivating the digital noise reduction system in the hearing aid.

Perceptual Quality Ratings

Quality ratings for the speech output through the hearing aid was done only for the participants in Group B1 and Group B2. Quality ratings were obtained in aided conditions with DNR activated and deactivated in order to answer the research question of whether there is any sound quality difference seen between the activated and deactivated DNR signal processing.

The participants in Group B were asked to rate the hearing aid in terms of quality of speech output, at 0 dB SNR when the DNR was activated and deactivated. For this, a recorded Kannada passage on the CD was routed to the loudspeaker through auxiliary input of the audiometer. The presentation level was at 45 dB HL, and white noise was also routed through the same loudspeaker such that the SNR was 0 dB.

The participants were instructed to listen carefully to the recorded paragraph which was presented and to rate on four parameters of quality. The participants were instructed to rate the quality on a 10-point rating scale in terms of loudness, clearness, naturalness and overall impression. For loudness, a rating of 9 was given when speech output through the hearing aid is sufficiently loud. In contrast, 0 was given if the speech was very loud /faint. For clearness, a rating scale of 9 was given when the speech was clear and distinct; whereas for blurred and distorted speech, the rating was 0. For naturalness, a higher rating was given when the speech sounded as if there was no hearing aid, i.e., natural. The overall impression is the output of speech with little distortion, giving rise to speech that was very similar that in quiet condition.

For this, the participants were asked to rate on a 10-point rating scale, where 0 is very poor and 9 is excellent. The rating of speech was done while listening to a recorded passage, through the hearing aid with DNR being activated and deactivated, only for participants in Group B1 and Group B2.

To assess the test re-test reliability of perceptual quality rating, the Group B participants were called to attend another session. Only five participants out of seven, in each Group of B attended the second session. The same instructions were given in the second session also. The gap between the two sessions was not less than 6 hours or more than one day.

Phase III: Electrophysiological Testing to record the auditory evoked late latency responses (ALLR

For each participant, a new recording session was created by entering and saving the details of patient's demographic data in the Bio-Logic Navigator Pro AEP system. The AEP system was calibrated to give a 65 dB SPL output of /da/ stimulus from a distance of 1 meter at 45^0 Azimuth. The white noise was also calibrated to give same output, such that 0 dB SNR was achieved.

The skin surface at two mastoids (M1, M2) and vertex (Cz) were cleaned with a skin preparation gel with a mild abrasive to obtain required impedance. It was ensured that the impedance at each electrode site was less than 5 kΩ and the inter-electrode difference in impedance was less than 2 kΩ. Silver chloride cup electrodes were used to record the responses and were placed in vertical montage. While recording ALLR, the non-inverting electrode (+) was placed on the vertex (Cz), the ground electrode was on mastoid of the non-test ear and the inverting electrode (-) on the mastoid of the test ear (M1 or M2). The participants were instructed to sit comfortably on a reclining chair and relax during the testing and they were asked to watch a muted movie played from a battery operated laptop. They were also instructed to ignore the stimulus and restrict the movement of head, neck and eye during testing.

The recorded natural /da/ stimulus was given through the loudspeaker, connected to Biologic Navigator Pro EP system, which was located at $45^{\,0}$ Azimuth and a distance of 1 meter from test ear. The non-test ear was given a 55 dB HL noise from the portable audiometer, in order to avoid its participation. To record ALLR in noise condition, white noise was routed to the same loud speaker at 0 dB SNR. The ALLR recording was initiated once a stable EEG was obtained. The stimulus and the recording parameters of speech evoked ALLR are given in the Table 1. The recording was done twice in each test condition to check for the replicability of the ALLR and weighted average of two recordings was taken.

The same procedure was followed for participants in Group B1 and Group B2 under two aided conditions. In the first aided condition, the ALLRs were recorded in the presence of noise at 0 dB SNR, by deactivating the digital noise reduction. In the second condition, ALLRs were recorded again in noise condition by activating the DNR in the hearing aid. Thus, the effect of DNR signal processing on the ALLR peaks was studied by comparing the two aided conditions.

Analysis of ALLR

The latency of the wave P1, N1 and P2 and amplitude of N1-P2 complex, in the two recordings were identified and marked visually by two experienced audiologists. The latencies of the peaks were tabulated for P1, N1 and P2. The peak-to-peak amplitude of N1-P2 was measured and tabulated. The latencies of components of ALLR (P1, N1 and P2) were marked at the center

Table 1: Table 1: Stimulus and acquisition parameters for recording of ALLR

Stimulus parameters	
Stimulus	Natural /da/
Intensity of stimulus	65 dB SPL
Transducer	Loud speaker at 45^0 azimuth, 1m
Mode of presentation	Monoaural
Number of samples	300
Stimulus polarity	Alternating
Repetition rate	1.1/sec
Ipsilateral masking	White noise (0 dB SNR)
Acquisition Parameters	
Filter setting	1-30 Hz
Notch filter	Off
Analysis window	-100 to +446 ms
No. of channel	Single channel
Amplification	50,000
Artifact rejection	75μV
Electrode Montage	
Non-inverting	Vertex (Cz)
Inverting	Test ear mastoid (A1/A2)
Ground	Non-test ear mastoid (A1/A2)

of the peak, if the peak was broader and if the peak was broader with unequal amplitude then the one with greater amplitude was marked.

In addition, the audiologists were also asked to rate the morphology of the waveforms, under the DNR activated and deactivated conditions, on a 5-point rating scale. Where 0 was used for no response, 1 for poor morphology, 2 for moderate morphology, 3 indicated good morphology whereas 4 for excellent morphology. The average of the ratings given by the two audiologists were calculated and tabulated.

For each participant in Group A, the SRS and ALLRs were obtained in quiet condition and with noise at 0 dB SNR. For each participant in Group B, the SRS and ALLRs were obtained under three aided conditions i.e., in quiet and with noise at 0 dB SNR, with DNR being activated and deactivated. In addition, for each participant in Group B, perceptual quality rating of speech output were obtained for four parameters when listening to speech through hearing aid under two conditions, when DNR was activated and in deactivated condition.

Results

All the statistical tests were performed using Statistical Package for Social Science software (version 16.0). The Shapiro-Wilk's test of normality was used to assess whether the sampling distribution between means was normal (Howell, 2008). Normality needs to be checked for each of the independent variables for each of the sample groups. The results showed that, most of the parameters were normally distributed, thus parametric tests were administered. However, the SRS data for Group A did not follow the normal distribution.

Effect of Noise (0 dB SNR) on SRS

Descriptive statistics was done on the SRS obtained in quiet condition and in noise at 0 dB SNR for Group A (N=10), Group B1 (N=7) and Group B2 (N=7) to compute the mean and standard deviation. The results are outlined in Table 2.

Table 2: Mean and standard deviation (in brackets) values of SRS (Max score: 25) obtained in quiet condition and in noise at 0 dB SNR, in the three groups

	Total correct scores		
Conditions	Group A	Group B1	Group B2
In quiet	24.40 (0.96)	20.86 (0.90)	19.86 (0.69)
In noise	20.10 (0.56)	17.00 (1.15)	16.14 (1.06)

From the Table 2, it can be inferred that as expected, the mean SRS in quiet is greater than the scores obtained in noise condition for all the three groups. The results of two-way repeated measures ANOVA indicated that there was no significant interaction between conditions and the groups [$F(2, 21) = 1.034$, $p>0.05$]. However, there was a significant main effect of condition [$F(1, 21) = 468.26$, $p<0.05$] and the group [$F(2, 21) = 70.837$, $p<0.05$].

The speech recognition scores of words are decreased by the addition of white noise, in all the three groups. Further, to investigate the degree to which these three groups are being affected by the white noise, the difference between the speech recognition scores in quiet and noise were calculated. Since the SRS obtained in quiet were not comparable across the groups, the mean of the difference in SRS (in %) were used to find out the impact of noise. These reductions in mean SRS obtained in percentage were greater for Group B2, while it is least for Group A.

Effect of Noise (0 dB SNR) on ALLR

The mean and standard deviation were obtained for latencies of P1, N1 and P2 and amplitude of N1-P2 complex, under quiet and noise condition, using descriptive statistics. The results showed that latencies of P1, N1 and P2 were significantly prolonged and there was reduction of N1-P2 amplitude in the presence of noise at 0 dB SNR in all the three groups. The effect the noise on P1 latency is similar across all the three groups. However, the latency of N1 obtained in quiet condition for individuals with hearing loss (Group B1 & Group B2) was significantly prolonged than participants with normal hearing (Group A). In addition, the P2 latency in individuals with moderate hearing loss (Group B2) is prolonged to a greater extent compared to individuals with mild hearing loss (Group B1). Therefore, this suggested that the effect of noise increases with increase in the degree of hearing loss.

Table 3: Mean and standard deviation (in brackets) of the difference values of SRS, across three groups

	Group A	Group B1	Group B2
SRS (quiet) - SRS (Noise)	4.30 (0.94)	3.86 (0.69)	3.72 (0.95)
Reduction in mean SRS (in percent)	17.62	18.50	18.73

Effect of DNR on SRS

Descriptive statistics was done to obtain the mean and standard deviation for the two aided conditions i.e., when DNR was activated and when DNR was deactivated. A look into the mean values in Table 4, indicates that the SRS of words in 'DNR activated' condition are greater than in 'DNR deactivated' condition. Hence, there is slight improvement in the speech recognition scores on activation of DNR.

Table 4: Mean and standard deviation values (in brackets) of SRS (Max score: 25) obtained in quiet and under noise at 0 dB SNR, with DNR activated and deactivated condition, across Group B1 and Group B2

	Total correct scores	
Condition	Group B1	Group B2
Quiet	20.86(0.90)	19.86(0.69)
DNR deactivated	17.00(1.15)	16.14(1.06)
DNR activated	17.86(0.90)	17.29(1.38)

Effect of DNR on ALLR

Descriptive analysis was done to obtain mean and standard deviation of P1, N1 and P2 latencies and amplitude of N1-P2 complex obtained in quiet and noise conditions, for the two groups of participants with hearing loss (Table 5). The mean latencies of P1, N1 and P2 across the two conditions shows that, the latencies recorded under DNR activation were significantly shorter in relation to the latencies obtained under DNR deactivated condition. But the amplitude of N1-P2 complex was not different across the two conditions (DNR activated and deactivated).

Table 5: Mean, standard deviation (SD) and p value of P1, N1 and P2 latencies and amplitude of N1-P2 complex, across two groups (Group B1 and Group B2), under DNR deactivated and activated condition

Components of LLR	Conditions/ p	Group B1	Group B2
P1 (in ms)	DNR deactivated	89.60 (8.83)	85.62 (6.67)
	DNR activated	81.87 (7.61)	81.19 (5.62)
	p	0.02*	0.01*
N1 (in ms)	DNR deactivated	140.98 (14.03)	135.51 (8.05)
	DNR activated	134.71 (15.77)	130.48 (8.57)
	p	0.08*	0.00*
P2 (in ms)	DNR deactivated	226.88 (11.93)	217.02 (10.56)
	DNR activated	219.81 (11.84)	206.46 (9.41)
	p	0.02*	0.01*
N1-P2 Amplitude(μV)	DNR deactivated	3.13 (0.44)	3.039 (0.83)
	DNR activated	4.03 (0.85)	4.264 (0.90)
	p	0.683	1.00

*Note: * indicates significant difference at 0.05 significance level*

Effect of DNR on perceptual quality ratings

The mean and standard deviation for four perceptual parameters under two test conditions (DNR deactivated & activated) for two groups was computed using descriptive statistics. Table 6 gives the mean and standard deviation values of the quality ratings on four perceptual parameters with DNR deactivated and activated condition.

It can be noted from Table 6 that the mean values of quality ratings is greater when the DNR was activated compared to when the DNR was deactivated, on all the four parameters for both Group B1 and Group B2. The mean values show that in DNR deactivted condition, the participants with mild hearing loss (Group B1) rated the quality under DNR deactivated to be significantly better than those with moderate hearing loss (Group B2), except for the 'Clarity' pararmeter. For the 'Clarity' parameter, the two groups (Group B1 & Group B2) did not signficantly differ in their ratings. Also, these groups were not significantly different when the DNR was activated.

This result suggests that the difference in the ratings obtained in DNR activated and deactivated condition is greater for Group B2 than Group B1, as Group B1 participants rated the 'Loudness', 'Clarity' and 'Overall impression' higher in DNR deactivated condition. In other words, the annoyance caused by the noise under the DNR deactivated condition is less disturbing for participants in Group B1 and hence the ratings given are more favourable; and vice versa for Group B2 participants. Therefore, it can be inferred that DNR implementation in hearing aid is more benificial for participants with moderate hearing loss (Group B2) than for those with mild hearing loss (Group B1).

Mann-Whitney U test was performed to compare the differences in quality ratings between the two groups of participants (Group B1 & Group B2) on a 10-point rating scale. This was done since the data was ordinal. The results showed significant difference in ratings of the perceptual parameters namely 'Loudness', 'Naturalness' and 'Overall impression' when DNR was deactivated between two groups (Group B1 & Group B2). Thus, the results of perceptual quality ratings are disscussed separately for Group B1 and Group B2. Also, these groups were not significantly different when the DNR was activated.

Wilcoxon Signed-Rank test was done for the pair-wise comparison of DNR activated and deactivated conditions, to evaluate the significance of difference (if any) between the two conditions, for Group B1 and Group B2. For Group B1 participants, the results showed that there was a significant difference ($p<0.05$) between the ratings of two perceptual quality parameters (Loudness & Clarity) obtained in DNR deactivated and DNR activated conditions. Although there was a difference between the ratings obtained in DNR activated and deactivated conditions, for the parameters 'Naturalness' and 'Overall impression', they are not statistically significant.

For Group B2 participants, the results revealed that there was statisically significant differences ($p<0.05$) seen between the quality ratings of all the four perceptual parameters under DNR activated condition and DNR deactivated condition. The 'Overall impression'

Table 6: Table 6: Mean and standard deviation (in brackets) values of four perceptual parameters of quality obtained with DNR activated and deactivated conditions, across the two groups

		Rating on a 10 - point scale (0-very poor, 9-excellent)	
Groups	Parameters of quality	DNR deactivated	DNR activated
Group B1	Loudness	6.71 (0.75)	8.00 (0.57)
	Clarity	6.86 (0.69)	8.57 (0.53)
	Naturalness	6.86 (0.90)	7.57 (0.53)
	Overall impression	7.00 (1.00)	8.14 (0.60)
Group B2	Loudness	5.57 (0.53)	7.57 (0.97)
	Clarity	5.71 (0.75)	7.86 (0.37)
	Naturalness	6.14 (1.06)	7.71 (0.48)
	Overall impression	6.43 (0.97)	8.57 (0.53)

of the quality through the hearing aid in background noise condition was rated the best compared to all other parameters, when DNR was activated. Comparable ratings were obtained for all other perceptual parameters, when the DNR was activated. It must be noted that the participants were asked to rate '9' (highest score) if the loudness of the speech was comfortable level; in contrast '0' was given when the signal was faint or too loud. Thus, the DNR signal processing signficantly improved the loudness of speech such that it is comfortable.

Test Retest Reliability

Test re-test reliability on quality rating was analyzed using Cronbach alpha, the intra-class correlation statistics for the ratings obtained in the two sessions. Cronbach's alpha is a coefficient of reliability and it normally ranges from 0 to 1, where 1 indicates excellent internal consistency.

The Cronbach Alpha was obtained across two sessions for four perceptual quality paramenters, under DNR activated and deactivated conditions. In DNR deactivated condition, the reliability between the two sessions for all the four parameters of quality ranged from acceptable to excellent reliability (Table 7).

Under the DNR activated condition, the tests re-test reliability for the three perceptual parameters viz., 'Loudness', 'Clarity' and 'Overall impression' revealed that there was a good to excellent reliability between the two sessions (Table 8).

Correlation

Spearman rank correlation co-efficient was obtained between N1-P2 amplitude and SRS, and between SRS and morphology ratings. The results indicated that, the no correlation was obtained between N1-P2 amplitude and SRS obtained in quiet and noise conditions in all the three Groups. The average of the morphology ratings given by two audiologists was tabulated for the statistical analysis. The results showed that there is no correlation obtained between SRS and morphology ratings in both activated and deactivated DNR conditions.

Discussion

The present results suggest that the speech recognition scores of words are significantly reduced in the presence of white noise. The above results are in accordance with the finding of studies in literature, which report that the speech recognition scores in the presence of noise are reduced when compared to that obtained under quiet condition (Keith & Talis, 1972; Carhart, Tillman, & Greetis, 1969; Danhauer, Doyle, & Lucks, 1985). These studies also report that the individuals with hearing loss

Table 7: Cronbach Alpha (A) value and reliability when the DNR was deactivated for four quality parameters for Group B

Perceptual parameters	A	Reliability
Loudness	0.8	Good
Clarity	0.9	Excellent
Naturalness	0.7	Acceptable
Overall impression	0.8	Good

Table 8: Cronbach Alpha (A) value and reliability when the DNR was activated for four quality parameters for Group B

Perceptual parameters	A	Reliability
Loudness	0.8	Good
Clarity	0.9	Excellent
Naturalness	-	-
Overall impression	0.8	Good

are more susceptible to background noise than individuals with normal hearing.

The result of the present study is in consonance with the findings reported in literature, which revealed that both participants with hearing loss (Group B1 & B2) are affected by noise to a greater degree than compared to individuals with normal hearing. This could be attributed to the reduced frequency selectivity and excessive upward spread of masking in individuals with hearing loss (Martin & Pickett, 1970; Trees & Turner, 1986).

Auditory processing of natural /da/ stimuli, at the cortical level is negatively affected by the presence of white noise, as indicated by smaller amplitude (N1-P2 complex) and increased latencies for ALLR components (P1, N1, & P2). These findings are in agreement with Martin and Stapells (2005), who investigated the effect of background noise on CAEPs in individuals with normal hearing. They used /ba/ and /da/ speech sounds to elicit the responses and they concluded that the latencies were significantly prolonged in the presence of noise compared to that in quiet condition. The reason for the prolonged latencies in the presence of noise, could be due to pronounced disruption of the timing features in cortical processing, when encoding rapidly presented acoustic signal that have been masked by noise (Wible, Nicol, & Kraus, 2004; Chandra & Barman, 2009).

Further, there was a reduction of N1-P2 amplitude in the presence of noise at 0 dB SNR in all the three groups. These results are in accordance with the findings reported by Martin and Stapells (2005); and Chandra and Barman (2009). They investigated the effect of noise on CAEPs by using /ba/ and /da/ speech stimulus. Their results indicated that the amplitude of N1-P2 reduced significantly in the presence of noise. Since ALLR is an exogenous potential, the components of ALLR namely P1, N1 and P2 depend on the characteristics of the stimulus. Hence, the presence of noise decreases the audibility of the stimulus leading to a reduction in N1-P2 amplitude and prolongation of latencies (Martin & Stapells, 2005; Chandra & Barman, 2009).

The participants with hearing loss did not perform equivalent to those with normal hearing as showed by ALLR latencies in quiet condition, even after providing appropriate amplification. It must be noted that the ALLRs were recorded in aided condition for Group B whereas unaided condition for Group A. The prolongation of latencies of P1, N1 and P2 in individuals with hearing loss (Group B1 & B2) could be due to the physiological changes such as damaged hair cells and auditory nerve fibers which result in elevated thresholds and broadened tuning curves that may affect the place and timing cues that are encoded throughout the auditory system. Further, the prolongation of latencies may also be influenced by the delay in the processing of the stimuli through the hearing aid. Therefore, in addition to hearing aid processing, damaged mechanisms in the peripheral auditory system might probably modify the signal before it reaches the brain (Souza & Tremblay, 2006). Hence, these individuals are not performing similar to individuals with normal hearing, under quiet condition.

Furthermore, the results showed that the activation of DNR in the hearing aid led to slight improvement in the speech recognition scores. However, the studies reported in literature (Boymans, Dreschler, Schoneveld, & Verschuure, 1999; Alcantara et al, 2003), indicated that the scores were similar across the activated and deactivated conditions. The inconsistency between the studies could be attributed to the speed of gain reduction, how fast the DNR is capable of reducing the noise and also the magnitude of gain reduction, degree to which noise suppression occurs. Further, differences in variables such as the type of competing signal and type of test stimuli used may play a role. The present study used the DNR capable of suppressing noise to moderate degree, competing signal was white noise and also words as test stimuli. On the other hand, Alcantara et al, (2003) used four different types of competing signal as well as low redundancy sentences as test stimuli. Whereas, in Boymans et al's (1999) study, modulation based DNR was used. The hearing aid used in the present study had frequency- based DNR algorithm (Figure 1).

The activation of DNR has reduced the prolongation of latencies in comparison to the latencies obtained under deactivated DNR condition. As reported in literature, the morphology of P1, N1 and P2 latencies is driven by the signal-to-noise ratio. As the SNR increases, the latencies of P1, N1 and P2 reduce (Billings et al, 2009). Since the activation of DNR reduces further deterioration of the latencies caused by the noise, the DNR

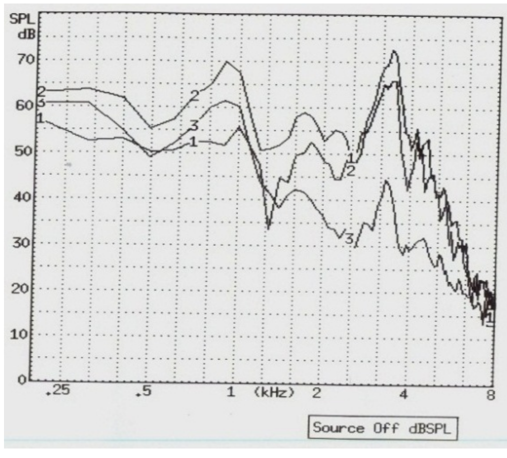


Figure 1: The electroacoustic measurements obtained for /da/ in quiet (curve 3), /da/ in DNR deactivated condition (curve 2) and /da/ in DNR activated condition (curve 1).

might probably be enhancing the SNR. However, this improvement obtained in the latencies under DNR activation was not similar to the latencies obtained in quiet condition.

The electroacoustic coupler measurements were done to investigate the gain changes across frequencies under quiet condition and under noise with DNR activated and deactivated conditions. Natural /da/ stimulus was given in quiet and noise. This was picked up by the hearing aid with / without DNR being activated. The output from the hearing aid was measured by Fonix 7000.

As shown in Figure 1, the curve 1 where the /da/ is given in noise with DNR activated had differential gain reduction for low frequencies and high frequencies. The maximum gain reduction was seen at the low frequencies when compared to high frequencies. Also the curve 1 (DNR activated) is not equivalent to the gain curve obtained under quiet condition (curve 3). Thus, the distortions of the stimuli caused by the activation of DNR which is evident through the acoustic measure could be attributed to the reduction in speech recognition scores (SRS) and delay in ALLR latencies seen under DNR activated condition, in comparison to quiet condition, in participants with hearing loss (Group B1 & B2).

For the perceptual quality ratings in Group B1 participants, the 'Naturalness' and 'Overall impression' of the speech with and without the DNR activation were not statistically different. This could be attributed to the milder form of hearing loss and also due to lack of acclimatization to the aided speech, as all the participants were naive hearing aid users (Ovegard, Lundberg, Hagerman, Gabrielsson, Bengtsson & Brandstrom, 1997).

DNR activation improved the sound quality of speech in terms of 'Loudness', 'Clarity', 'Naturalness' and 'Overall impression' for Group B2 participants. This result is in accordance with the studies reported in literature, which report that implementation of DNR in hearing aid leads to improvement in sound quality (Boymans et al, 1999; Ricketts & Hornsby, 2005).

The results of the study also indicated that there is no relationship between SRS and amplitude of ALLR. These results are in agreement with the study done by Chandra and Barman (2009). They attributed the lack of correlation between speech recognition scores and ALLR to the wide variability of latencies and amplitude of ALLR across the subjects. In addition, the components of ALLR are affected by a number of factors such as background EEG, impedance between the electrodes, sleep or drowsiness state etc., which might have led to poor correlation (Chandra & Barman, 2009). Thus, speech recognition scores will not only depend on generators of ALLR, but also on other factors.

Conclusions

From the study, it can be concluded that implementing a frequency-based DNR will lead to slight improvement in scores of speech recognition, reduces the prolongation of ALLR peaks and also improves the sound quality of speech in the presence of noise. However, DNR is more beneficial to the participants with higher degree of hearing loss.

The results of the ALLR in the present study help in understanding how the signal is encoded at the cortical level, in the presence of noise in individual with normal hearing and hearing loss. Speech perception inferred through electrophysiological measures such as ALLR has two advantages. ALLR as an objective measure does not require the active participation of subjects; and also the speech stimulus used to record ALLR is not language specific and hence can be used for wide range of population. The DNR has improved the sound quality in the presence of noise, thus there is greater chance of using the hearing aid more often in day-to-day life, than rejecting the hearing aid.

References

Alcantara, J. L., Moore, B. C., Kuhnel, V., & Launer, S. (2003). Evaluation of the noise reduction system in a commercial digital hearing aid. *International Journal Audiology, 42*, 34-42.

Billings, C., Tremblay, K., Steckera, G., & Tolina, M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research, 254*(1-2), 15-24.

Boymans, M., & Dreschler, W. A. (2000). Field trials using a digital hearing aid with active noise reduction and dual-microphone directionality. *Audiology, 39*(5), 260-268.

Boymans, M., Dreschler, W. A., Schoneveld, P., &Verschure, H. (1999). Clinical value of a full-digital in-the-ear hearing instrument. *Audiology, 38*, 99-108.

Bray, V., & Nilsson, M. (2001). Additive SNR benefits of signal processing feature in a directional hearing aid. *Hearing Review, 8*, 48-51.

Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America, 45,* 694-703.

Chandra, R., & Barman, A., (2009). *Relationship between ALLR and SIS in individuals with auditory dys-sychrony.* Master's dissertation, University of Mysore, India.

Danhauer. J. L., Dolye, P. C., & Luks, L. (1985). Effects of noise on NST and NU-6 stimuli. *Ear and Hearing, 6*, 266-269.

Gabrielsson, A., Schenkman, B. N., & Hagerman, B. (1988).The effects of frequency responses on sound quality judgements and speech intelligibility. *Journal of Speech Hearing Research, 31*, 166-177.

Gelfand, S. A., Piper, N., & Silman S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *Journal of the Acoustical Society of America, 80,* 1589-1598.

Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., & Muchnik, C. (2006). Identification of syllables in noise: electrophysiological and behavioral correlates. *Journal of the Acoustical society of America, 120*(2), 926-933.

Keith, R. W., & Talis, H. P. (1972).The effects of white noise on PB scores of normal and hearing-impaired listeners.*Audiology, 11*(3), 177-186.

Kochkin, S. (2005). Consumers satisfaction with hearing instruments in the digital age. *The Hearing Journal, 58*(9), 30-43.

Martin, B. A., & Stapells, D. R. (2005). Effects of low-pass noise masking on auditory event-related potentials to speech. *Ear and Hearing, 26*(2), 195-213.

Mueller, H. G., Weber, J., & Hornsby, B.W. (2006).The effects of digital noise reduction on the acceptance of background noise. *Trends in Amplification, 10*, 83-93.

Nabelek, A. K. (1988). Identification of vowels in quiet, noise, and reverberation: Relationships with age and hearing loss. *Journal of the Acoustical Society of America, 84,* 476-484.

Narne, V. K., & Vanaja, C. S. (2008).Speech identification and cortical potentials in individuals with auditory neuropathy. *Behavioral and Brain Functions, 15*(4), 123-130.

Nordrum, S., Erler, S., Garstecki, D., & Dhar, S. (2006). Comparison of performance on Hearing in Noise test using directional microphone and digital noise reduction algorithms. *American Journal of Audiology, 15*, 81-91.

Ovegard, A., Lundberg, G., Hagerman, B., Gabrielsson, A., Bengtsson, M., & Brändström, U. (1997).Sound Quality Judgement during Acclimatization of Hearing Aid. *Scandinavian Audiology, 26*(1), 43-51.

Pittman, A. L., & Stelmachowicz, P. G. (2003).Hearing loss in children and adults: Audiometric configurations, asymmetry, and progression. *Ear and Hearing, 20*(3), 198-205.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *Journal of the Acoustical Society of America, 63*, 533-549.

Ricketts, T., & Hornsby, B. (2005). Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. *Journal of the American Academy of Audiology, 16*, 270-277.

Scherer, M. J., & Frisina, D. R. (1998). Characteristics associated with marginal hearing loss and subjective well-being among a sample of older adults. *Journal of Rehabilitation Research and Development. 35,* 420-426.

Souza, P. E. & Tremblay, K. L. (2006).New Perspectives on Assessing Amplification Effects. *Trends in Amplification, 10*(3), 119-142.

Trees, D. E., & Turner, C. W. (1986). Spread of masking in normal subjects and in subjects with high frequency hearing loss. *Audiology, 25,* 70-83.

Vanaja, C. S. & Jayaram, M. (2006). *Sensitivity and specificity of different audiological tests in differential diagnosis of auditory disorders.* Developed in Department of Audiology, AIISH, Mysore. India.

Weinstein, B. E, & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *Journal of Speech and Hearing Research, 25*(4), 593-599.

Whiting, K. A., Martin, B. A., & Stapells, D. R. (1998). The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. *Ear and Hearing, 19*(3), 218-231.

Wible, B., Nicol, T., & Kraus, N. (2004). A typical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychological, 67*(3), 299-317.

Yathiraj, A., & Vijayalakshmi, C. S., (2005) *Phonemically balanced word test in Kannada*. Developed in Department of Audiology, AIISH, Mysore. India.

Assessment of Different Vestibular Pathways in Individuals with Dizziness

[1]Saravanan P. & [2]Sinha, S. K.

Abstract

Assessment of vestibular system is complex since vestibular system involves multiple numbers of structures. Hence for the assessment of the complete vestibular system multiple numbers of tests are required. Two groups of subjects participated in the present study, Clinical group consists of 15 participants in 3 sub groups with 2 groups having an otological diagnosis and the third group having no otological or neurological diagnosis for vertigo. Sub group A consisted of 8 participants (4 males & 4 females) with Meniere's disease. 7 of the participants had unilateral Meniere's disease and one participant had bilateral Meniere's disease. Sub group B consisted of 5 participants (3 males & 2 females) with unilateral vestibular neuritis. Sub group C consisted of 2 participants (1 male & 1 female) with idiopathic condition. Control group consisted of fifteen individuals (11 males & 4 females) without complaint of dizziness. In control group, the mean latencies of oVEMPs peaks were significantly shorter than mean latencies of cVEMPs peaks. Amplitude of cVEMP peak complex was significantly greater compared to the amplitude of oVEMPs peaks. Also in the contralateral ear of the some participants with Meniere's disease VEMPs results were abnormal. In the sub group B, depending on the involvement of the superior and inferior vestibular nerves the abnormal results were found in ENG, cVEMPs and oVEMPs. No significant association was seen between any of the two tests in both sub group A and B. In the sub group C, the ENG, cVEMPs and oVEMPs results clearly indicate involvement of multiple of anatomical structures.

Keywords: *Dizziness, vestibular functions, VEMP, ENG*

Introduction

Balance may be defined as the capacity to maintain posture and spatial orientation at rest and during movement. The sensory inputs for maintenance of equilibrium comes from three main systems i.e., visual, proprioceptive and vestibular system. Disturbances in any of these systems results in perception of disequilibrium.

The vestibular system is the organ of balance, helps to maintain a balanced position in three-dimensional space. Vestibular stimulation results in three types of reflex responses: vestibulo-ocular, vestibulo-spinal, and vestibulo-collic reflexes. The vestibulo-ocular reflexes help to maintain gaze on a stationary object while the head or body is in motion. Two types of the vestibulo-ocular reflex system are the 'semicircular canal ocular reflex, and the 'otolith ocular reflex' (Bronstein & Gresty, 1991). Vestibulo-spinal reflex is primarily responsible for control of tone in skeletal muscles of the trunk and extremities. Vestibulo-collic reflexes are thought to act on neck muscles in order to stabilize the head, especially during unpredictable movements (Schubert & Shepard, 2008). Any disturbance to these reflexes leads to vestibular dysfunction. Vestibular dysfunction can be peripheral and/or central. Peripheral vestibular dysfunction refers to dysfunction of vestibular end organs or vestibular nerve. Central vestibular dysfunction involves dysfunction of vestibular nuclei, cerebellum, or the occulomotor, vestibulospinal, and proprioceptive pathways.

Assessment of vestibular function includes battery of clinical, electrophysiological and questionnaire based tests. The audiological tests which are administered to assess the vestibular functions are detailed case history, puretone audiometry, immitance measurment, speech audiometry, otoacoustic emissions and auditory brainstem responses, electronystagmography (ENG) and vestibular evoked myogenic potentials (VEMPs).

Current electrophysiological evaluation of the vestibular system, such as ENG, does not assess all functional structures and pathways. ENG battery only assesses lateral semicircular canals and the superior vestibular nerve. By adding cervical VEMPs (cVEMPs) measurements, it is possible to identify any dysfunction in the saccule and/or inferior vestibular nerve.

Another variant of VEMPs are the ocular vestibular evoked myogenic potentials (oVEMPs). oVEMPs are likely to be produced by synchronous activity in the extraocular muscles, i.e., myogenic potentials (Rosengren, Todd & Colebatch, 2005). oVEMPs responses mainly assess the function of otolith organs and superior vestibular nerve.

Studies have reported that the clinical application of ENG (Bergman & Stahle, 1967; Thomas & Harrison, 1971; Wennmo & Pyykko, 1982), cVEMPs (Akkuzu, Akkuzu & Ozluoglu, 2006; Boleas-Aguirre, Sanchez-Ferrandiz, Artieda & Perez, 2007; Murofushi, Shimizu, Takegoshi & Cheng, 2001) and oVEMPs (Chiarovano, Zamith, Vidal & Waele, 2011; Huang, Wang & Young, 2012; Murofushi, Nakahara, Yoshimura & Tsuda, 2011) in various vestibular dysfunction.

cVEMPs are sensitive to disorders affecting saccule or

[1]Email: saravananpaud@gmail.com,
[2]Lecturer in Audiology, Email: sujitks5@gmail.com

the inferior vestibular nerves. cVEMPs in individuals with vestibular symptoms such as "objects spinning or turning around you" have been found to be absent or abnormal (Kumar, 2006). cVEMPs assess only vestibulo-collic reflex pathway whereas ENG assesses the integrity of semicircular canal ocular reflex pathway of the vestibular system. So, addition of ENG with VEMPs may provide valuable information for differential diagnosis in individuals presenting with vestibular dysfunction.

oVEMPs is the new kind of variation of VEMPs responses. Combining oVEMPs and cVEMPs provide complementary information about saccular and utricular otolithic function. Also there is dearth of information regarding oVEMPs recording in the clinical population. So there was need to study the oVEMPs in individuals with vestibular dysfunction.

Equivocal studies are available in the literature regarding the diagnostic significance of ENG in patients with vestibular dysfunction (Coats, 1970; Pflatz, 1984; Ojala, Vaheri & Juntenen, 1989; Colledge, Barr-Hamilton, Lewis, Sellar & Wilson, 1997; Shi, Yu, Niu & Lu, 1997; Bhansali & Honrubia, 1999; Bakr & Saleh, 2000). The equivocal findings might be due to the fact that, earlier tests for assessing the integrity of sacculo-collic pathway and otolith ocular reflex pathway were not available thus, there is need to study the integrity of different vestibular pathways (semicircular canal ocular reflex pathway, vestibulo-collic reflex pathway and otolith ocular reflex pathway) using ENG, cVEMPs and oVEMPs in differential diagnosis of vestibular dysfunction.

Equivocal findings also have been reported in association of ENG, cVEMPs and oVEMPs responses in individuals with various vestibular dysfunctions (Chiarovano et al., 2011; Jacobson et al., 2011; Murofushi et al., 2011). Hence there was need to study ENG, cVEMPs and oVEMPs in individuals with vestibular dysfunction. Combination of ENG, cVEMPs and oVEMPs may provide valuable information regarding the pathways involved in different vestibular dysfunction.

Aim of the study

The present study was aimed to assess different vestibular pathways (Semicircular canal ocular reflex pathway, Vestibulo-collic reflex pathway & Otolith-ocular reflex pathway) in individuals with vestibular dysfunction.

Method

Participants

Two groups of subjects participated in the present study, clinical group and control group. Clinical group consists of 15 participants in 3 sub groups with 2 groups having an otological diagnosis and the third group having no otological or neurological diagnosis for vertigo. Sub group consists of 8 participants (4 males & 4 females) with Meniere's disease participated in this study with age range between 18 to 55 years (Mean age = 37 years). 7 of the participants had unilateral Meniere's disease and one participant had bilateral Meniere's disease. Totally 9 ears with Meniere's disease were considered for this study. In sub group B, 5 participants (3 males & 2 females) with unilateral vestibular neuritis with the age range between 34 to 47 years (Mean age = 41 years) participated. In sub group C, 2 participants (1 male & 1 female) with idiopathic condition. The diagnosis of the participant could not be established. The age of the two participants was 34 and 50 years respectively. Fifteen individuals (11 males & 4 females) aged between 22 to 50 years with the mean age of 30 years were served as controls in the present study.

Instrumentation and Test Environment

Pure Tone Audiometry was done to confirm bilateral normal hearing sensitivity. Immittance audiometry was done to rule out middle ear abnormalities. Intelligent Hearing systems (IHS version 4.3.02) was used for recording auditory brainstem responses and air conducted click evoked cervical VEMPs. Biologic navigator Pro EP instrument with biologic insert was used for ocular VEMPs recording. All the audiological tests were conducted in the acoustically treated rooms and noise levels during the testing were within permissible limits (ANSI, 1991). ENG was conducted in a room with low ambient lighting.

Procedure

A detailed case history was taken for each participant prior to testing. It was followed by administration of the dizziness questionnaire. Pure tone audiometry and immittance evaluation was done for all the participants. Then the auditory brainstem responses (ABR) were recorded for both the ears to rule out any retro cochlear pathology. Two channel ABR recording was done for 100?sec click stimuli at 90 dBnHL with the rarefaction polarity. The repetition rate used was 11.1/sec and 90.1/sec. The responses were filtered between 100 Hz to 3000Hz.

During the cVEMPs recordings the participants were instructed to sit straight and turn their head to the opposite side of the ear in which stimulus was presented, so as to activate ipsilateral sternocleidomastoid (SCM) muscle, as it gives reliable and greater amplitude. cVEMPs was recorded using 500 Hz tone burst (2 cycles rise, 0 cycles plateau, and 2 cycles fall, Blackman weighting function) presented at a rate of 5.1/sec using rarefaction polarity. 500 Hz tone burst stimuli was used as the 500 Hz tone burst stimulus gives better amplitude of the cVEMPs (Kumar, Sinha, Bharti & Bar-

Table 1: Recording protocol for Electronystagmography

Band-pass filter	0.1 Hz to 30 Hz
Notch filter	On
Gain	Gain of the incoming signal will be adjusted in such a way that 10 mm deflection of recording pen represents 200 μv of corneoretinal potentials
No. of channels	1
Electrode placement	Non-inverting electrode (+): Outer canthus of the right eye Inverting electrode (-): Outer canthus of the left eye Ground electrode: Lower forehead

man, 2003). The stimuli were presented to the test ear at single intensity of 95 dBnHL using ER - 3A insert ear phones. The responses were recorded for 70 msec post stimulus period along with the 10 msec pre-stimulus period. The recorded responses were then amplified (X 5000) and band pass filtered between 30 to 1500 Hz. The responses were averaged totally for 200 stimuli. cVEMPs was recorded twice to ensure the replicability of the responses.

oVEMPs was recorded for all the participants with upper gaze direction. Participants were instructed to maintain the same upper gaze throughout the test run. Stimuli used to record oVEMPs were identical to stimuli used to record cVEMPs. 500 Hz tone burst (2 cycles rise, 0 cycles plateau, and 2 cycles fall, Blackman weighting function) presented at a rate of 5.1/sec using rarefaction polarity. The stimuli were presented monaurally at single intensity of 95 dBnHL using ER - 3A insert ear phones. 100 stimuli were used for response averaging. The response was analysed for 60 msec post stimulus period. A pre-stimulus period of 10 msec was utilised to record background electrical activity. The recorded electrical responses were amplified (X 5000) and band pass filtered between 1 Hz to 1000 Hz. oVEMPs responses were recorded twice in each ear to ensure replicability of the responses. ENG was recorded with the protocol as shown in Table-1.

Prior to the testing, ENG equipment will be calibrated for each participant. Calibration of the ENG instrument will be done for 10^o eye deviation and the input sensitivity of the instrument will be adjusted in such a way that every 10^o of eye movement corresponds to 10 mm movement of the recording paper and the 10 mm of paper will move in 1 sec. Two set of tests were administered using ENG system. The details of the test done are shown in Figure-1

Results and Discussion

Vestibular evoked myogenic potentials findings in Control group

Cervical - vestibular evoked myogenic potentials (cVEMPs)

cVEMPs responses could be recorded in all the participants in the control group. In cVEMPs, the latency of p13 and n23 peaks, amplitude of p13-n23 complex and amplitude asymmetry (between the two ears) were analyzed.

Descriptive statistics was done to find out mean and standard deviation for p13 and n23 latencies, amplitude of p13-n23 complex, and inter-ear amplitude asymmetry for p13-n23 complex. The descriptive results of latency, amplitude of p13-n23 complex and inter-ear amplitude asymmetry for p13-n23 complex for the control group is shown in Table 2. Bilateral cVEMPs recordings for one participant in the control group in response to 500 Hz tone burst stimuli presented at 95 dBnHL is

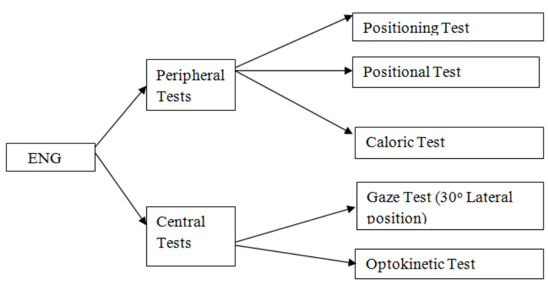


Figure 1: Flow chart of ENG test battery.

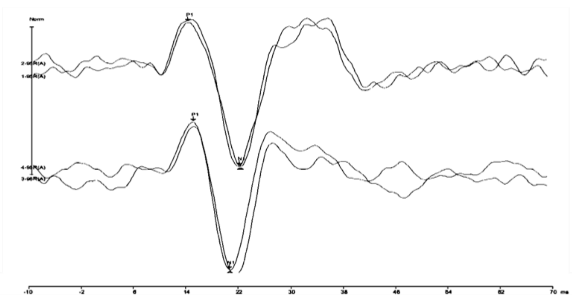


Figure 2: cVEMPs response of one control group participant.

Table 2: Mean and standard deviation (SD) values of latency and amplitude measures of cVEMPs in control group

Parameters	Mean	Standard deviation	Range
p13 latency (msec)	15.10	1.24	12.80 - 17.60
n23 latency (msec)	22.45	2.05	19.20 - 27.20
Amplitude of p13-n23 complex (μv)	40.39	12.66	19.78 - 87.97
p13- n23 complex inter-ear amplitude Asymmetry (%)	14.45	8.71	1.20 - 32.03

shown in Figure 2.

Ocular - vestibular evoked myogenic potentials

Latency of n1, p1 and n2, peak to peak amplitude of n1-p1 and p1-n2 complex, and inter-ear amplitude asymmetry were analyzed in oVEMPs. The oVEMPs could be recorded in all the participants of the control group. Mean and standard deviation of latency of n1, p1 and n2 are shown in Table 3 and amplitude measures are shown in Table 4. The bilateral oVEMPs recordings of one control group participant to 500 Hz tone burst stimuli presented at 95 dBnHL is shown in Figure 3.

Table 3: Mean and standard deviation (S.D) values of latency measures of oVEMPs in control group

Parameters	Mean	Standard deviation	Range
n1 latency (msec)	11.37	0.96	9.91 - 13.20
p1 latency (msec)	16.49	0.90	14.70 - 18.45
n2 latency (msec)	21.83	1.81	18.91 - 27.66

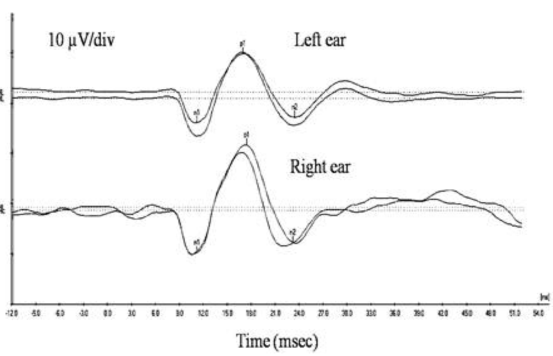


Figure 3: oVEMPs response of one control group participant.

Table 4: Mean and standard deviation (SD) values of amplitude measures of oVEMPs in control group

Parameters	Mean	Standard deviation	Range
Amplitude of n1-p1 complex (μv)	9.02	6.31	1.77 - 23.56
Amplitude of p1-n2 complex (μv)	8.16	5.62	1.46 - 20.54
n1-p1 complex inter-ear amplitude Asymmetry (%)	22.55	15.44	3.00 - 49.35
p1-n2 complex inter-ear amplitude Asymmetry (%)	20.73	12.51	0.68 - 44.49

Correlation between cVEMPs and oVEMPs responses in Control Group

To find out any significant correlation between the latency and amplitude of cVEMPs with latency and amplitude of oVEMPs, a Pearson correlation analysis was done. Pearson correlation analysis revealed a significant correlation between n23 latency of cVEMPs and n1 latency of oVEMPs ($r = 0.496$, $p<0.05$), between n23 latency of cVEMPs and p1 latency of oVEMPs ($r = 0.490$, $p<0.05$) and between n23 latency of cVEMPs and n2 latency of oVEMPs ($r = 0.490$, $p<0.05$). However, Pearson correlation analysis failed to show any correlation between the p13 latency of cVEMPs and p1 latency of oVEMPs ($r= 0.10$, $p>0.05$), p13 latency of cVEMPs and n1 latency of oVEMPs ($r= 0.15$, $p>0.05$), p13 latency of cVEMPs and n2 latency of oVEMPs ($r= 0.20$, $p>0.05$).

Amplitude of p13-n23 complex of cVEMPs did not show any significant correlation with amplitude of n1-p1 complex of oVEMPs ($r = 0.071$, $p>0.05$) and amplitude of p1-n2 complex ($r = 0.042$, $p>0.05$) of oVEMPs. Inter ear amplitude asymmetry of p13-n23 complex in cVEMPs was not significantly correlate with asymmetry of n1-p1 ($r = - 0.508$) or p1-n2 ($r = - 0.186$) complex of oVEMPs ($p>0.05$, Pearson correlation).

The latency of p13, n23 and the p13-n23 amplitude complex of cVEMPs and n1, p1 latency and n1-p1 amplitude complex of oVEMPs obtained in the present study are similar to earlier reports (Akin & Murnane 2001; Akin, Murnane & Medley 2003; Smulders et al., 2009; Chiarovano, et al., 2011; Murnane, Akin, Kelly & Byrd, 2011; Bohra, Sanju & Sinha, 2012). However, in the present study, an additional peak ‘n2’ in oVEMPs was observed consistently in all the subjects in the control group. The ‘n2’ peak has not been reported earlier and in the present study the latency of ‘n2’ was around 22 msec. It is hypothesized that the generators of the 'n2' peak also might be confined in the same anatomical structures from where the 'n1' and 'p1' peak is generated.

Latencies of oVEMPs is shorter compared to the cVEMPs responses and also there was no correlation between the latency of cVEMPs versus latency of oVEMPs. The differences in latencies between cVEMPs and oVEMPs might be due to the differences in length and nerve conduction velocity between vestibular ocular (Broussard et al, 1992) and vestibulo spinal pathways (Uchino et al, 2005) as shown in the animal studies. cVEMPs and oVEMPs responses are generated from different anatomical pathways stimulated by the air conduction stimulation (Chiarovano et al, 2011). cVEMPs responses mainly assess the function of saccule and inferior vestibular nerve, whereas, oVEMPs mainly assess the function of otolith organs and superior vestibular nerve. Pathway involved in the cVEMPs includes the saccular macula, inferior vestibular nerve, the lateral vestibular nucleus, the medial vestibulospinal tract, and the motor neurons of ipsilateral sternocleidomastoid (SCM) muscle (Halmagyi & Curthoys, 1999). The neuronal pathway for oVEMPs via the vestibulo-ocular reflex include, activation of the vestibular nerve and vestibular nucleus, medial longitudinal fasciculus, oculomotor nuclei, ocular nerves and to the contralateral extraocular muscles (Rosengren, et al.2010).

Amplitude of p13- n23 complex was significantly greater compared to the amplitude of n1-p1 and p1-n2 complex of oVEMPs. This is largely due to the differences in the muscle unit content between SCM and extraocular muscles (Park et al, 2010). The muscle thickness is more at the SCM compared to the extraoccular muscles and hence the tonic activation is more for the SCM compared to the extraoccular muscles (Park et al. 2010). Only a weak correlation was found between n23 latency of n1, p1 and n2. Amplitude measures of cVEMPs and oVEMPs did not correlate significantly. This may be due to large variability in amplitude measures of cVEMPs and oVEMPs.

Electronystagmography findings in control group

Central tests

*Optokinetic test:*In control group, all the participants had symmetrical optokinetic responses for visual target moving from right to midline and from left to midline.

Gaze test: In all the participants, there was no nystagmus for either in the left or right 30^{o} lateral position of gaze.

Peripheral tests

Positional test: Out of 15 subjects in the control group, 5 subjects showed some nystagmus beats in one or two head positions, but none of them had presence of nystagmus beats in more than two head positions. Maximum number of nystagmus beats recorded from subjects in the control group in any of the head position was 6 beats/30 seconds.

Table 5: Range of culmination frequency/30 seconds for all four caloric stimulation in control group

Caloric stimulation	Range of culmination frequency per 30 seconds
Right warm	22 - 59
Left warm	20 - 70
Right cold	21 - 51
Left cold	22 - 64

*Positioning test:*Out of 15 subjects in the control group only 3 subjects showed some nystagmus beats in one or two head positions, but none of them had presence of nystagmus beats in more than two head positions. Maximum number of nystagmus beats recorded in positioning test was 8 beats/ 30 seconds.

Caloric test: Bithermal caloric test was recorded from all the subjects in the control group. The culmination frequency was calculated for all the participants in the control group.

In control group, the range of culmination frequency of nystagmus in response to different caloric stimulation is shown in Table 5.

Claussen's butterfly chart was made from the culmination frequency obtained from the participants in the control group. Figure 4 shows a butterfly chart obtained from one of the participants in the control group.

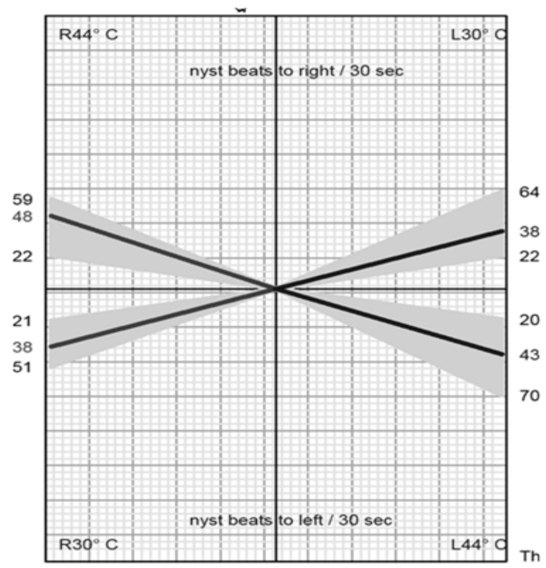


Figure 4: Results of caloric test for one participant in control group as shown in butterfly chart.

Table 6: Clinical information of 7 participants with unilateral Meniere's disease and one participant with bilateral Meniere's disease

				Air conduction thresholds (dBHL)						
S.No	Age (years)	Sex	Ear	250 Hz	500 Hz	1kHz	2kHz	4kHz	8kHz	ABR
1	50	F	Right	15	15	10	10	5	10	N
2	45	M	Right	35	30	35	40	35	40	N
3	26	M	Right	45	40	45	45	50	45	N
4	34	M	Right	55	50	40	45	40	45	N
5	$36^{\pm}$	F	Right	35	30	25	25	30	25	N
6	$36^{\pm}$	F	Left	45	40	40	35	30	35	N
7	18	F	Left	50	40	40	35	40	40	N
8	55	M	Right	40	35	30	25	25	25	N
9	32	F	Right	45	50	55	50	55	55	N

N - Normal, M-Male, F-Female, $\pm$ - Bilateral Meniere's disease, ABR - auditory brainstem response

In the gaze test all the participants of the control group showed no nystagmus in the left or right 30^0 lateral position of the gaze. This was similar to the findings as reported by Kirtane (2009).

Symmetrical optokinetic responses were recorded when the target moves right to midline or left to midline. Symmetrical optokinetic responses are reported in the healthy subjects (Kirtane, 2009). In the present study, none of the participants in the control group showed positional nystagmus with eyes closed condition in more than 2 positions. Isolated occurrences of positional nystagmus with the eyes closed condition in less than 3 positions is seen in some normal individuals (Kirtane, 2009). None of the participants had nystagmus beats in more than 2 head positions and during positioning test, similar findings are considered as normal (Kirtane, 2009).

It should be noted that from the results of caloric tests in the control group caloric responses in two ears and to warm and cold stimulation is not same. Warm stimulation had slightly more number of nystagmus beats compared to the cold stimulation in both the ears. Cold stimulation is reported to be stronger compared to the warm stimulation (Kirtane, 2009). The differences in the results might be due to methodology, type of irrigation and different instruments used in different studies.

Vestibular Findings in Individuals with Meniere's disease

Total 8 subjects (9 ears) in the Meniere's disease group were evaluated using the vestibular evoked myogenic potentials (cVEMPs and oVEMPs) and electronystagmography. 8 subjects had a unilateral Meniere's disease whereas one subject had a bilateral Meniere's disease. In the electronystagmography, both peripheral tests and central tests were administered to the participants. The details of the participants with Meniere's disease are given below in Table 6.

Vestibular evoked myogenic potentials

Cervical - vestibular evoked myogenic potentials

Out of 9 ears (8 participants) with Meniere's disease, 5 ears had absence of cVEMPs responses (55.55%), one ear had reduced peak to peak amplitude (11.11%) and 3 ears had normal responses (33.33%). Kruskal wallis test was performed to see whether the responses for 3 ears having normal cVEMPs responses were similar to the control group. Kruskal wallis test revealed no significant difference between cVEMPs responses latencies of p13 and n23, amplitude of p13-n23 between the control group and the Meniere's disease group ($p>0.05$)

Ocular - vestibular evoked myogenic potentials

In oVEMPs recordings, 7 affected ears showed absent responses (77.78%) and 2 ear showed present oVEMPs response (22.22%). The latency and amplitude of two ears with present oVEMPs was within the range of normal values.

Electronystagmography

Central tests

Optokinetic test: None of the participants in the Meniere's disease group showed any asymmetrical response in the optokinetic test.

Gaze test: All the participants showed no nystagmus in both gaze conditions.

Peripheral tests

Positional test: In 2 subjects with Meniere's disease, there were positional nystagmus present in two head positions, but the nystagmus beats were lesser than that obtained for the control group in any of the head positions in the present study.

*Positioning test :*In 3 subjects with Meniere's disease, nystagmus were present in two head positions in the positioning tests, but the nystagmus beats were lesser than that obtained for the control group in any of the head positions.

Caloric test: Out of 9 ears with Meniere's disease, ENG results showed hypo activity in 7 affected ear (77.78%), one ear showed hyper activity (11.11%) and one ear showed normal response (11.11%) to caloric stimulation.

Association of Caloric test, cVEMPs and oVEMPs results in subjects with Meniere's disease

Out of 9 ears with Meniere's disease, 8 ears (88.89%) showed abnormal results in caloric test, 7 ears (77.78%) showed abnormal or absent oVEMPs responses and 6 ears (66.67%) showed abnormal or absent cVEMPs responses.

In 66.67% (6 ears) individuals, there was association of caloric test and oVEMPs findings. 6 ears showed (66.67%) abnormal results in both caloric test and oVEMPs. None of the ear showed normal findings in both the tests. In 33.33% (3 ears) of individuals, there was dissociation of caloric test and oVEMPs responses. Two ear with normal oVEMPs responses showed abnormal caloric test findings, whereas, one ear with normal caloric test showed absent oVEMPs response.

In 55.56% (5 ears) individuals, there was association of oVEMPs and cVEMPs findings. In 4 ears (44.44%) with Meniere's disease, both oVEMPs and cVEMPs were absent. One ears (11.11%) showed normal response in both the test. In 44.44% (4 ears) of the individuals, there was a disassociation between cVEMPs and oVEMPs responses. 2 ears with normal cVEMPs responses showed absent oVEMPs responses. One ear with absent cVEMPs responses showed normal oVEMPs responses. One ear with reduced p13-n23 amplitude of cVEMPs showed absent oVEMPs response.

In 55.56 % (5 ears) individuals there was association of caloric test and cVEMPs findings. 5 ears showed (55.56%) abnormal results in both caloric test and oVEMPs. None of the ear showed normal findings in both the tests. In 44.44% (4 ears) of individuals, there was a disassociation between caloric test and cVEMPs responses. One ear with normal caloric test showed absent cVEMPs responses. 3 ears with abnormal caloric test findings showed normal cVEMPs responses.

To find out any significant association between the Calorics tests and cVEMPs, Caloric tests and oVEMPs, oVEMPs and cVEMPs responses, a Chi square test was done. Chi-square test revealed no significant association between caloric and oVEMPs findings ($p = 0.778$), caloric and cVEMPs findings ($p = 0.667$) and between cVEMPs and oVEMPs findings ($p = 0.583$).

In 55.56% (5 ears) individuals, there was association of oVEMPs and cVEMPs findings. In 4 ears (44.44%) with Meniere's disease, both oVEMPs and cVEMPs were absent. One ears (11.11%) showed normal response in both the test. In 44.44% (4 ears) of the individuals, there was a disassociation between cVEMPs and oVEMPs responses. 2 ears with normal cVEMPs responses showed absent oVEMPs responses. One ear with absent cVEMPs responses showed normal oVEMPs responses. One ear with reduced p13-n23 amplitude of cVEMPs showed absent oVEMPs response.

In 55.56 % (5 ears) individuals there was association of caloric test and cVEMPs findings. 5 ears showed (55.56%) abnormal results in both caloric test and oVEMPs. None of the ear showed normal findings in both the tests. In 44.44% (4 ears) of individuals, there was a disassociation between caloric test and cVEMPs responses. One ear with normal caloric test showed absent cVEMPs responses. 3 ears with abnormal caloric test findings showed normal cVEMPs responses.

To find out any significant association between the Calorics tests and cVEMPs, Caloric tests and oVEMPs, oVEMPs and cVEMPs responses, a Chi square test was done. Chi-square test revealed no significant association between caloric and oVEMPs findings ($p = 0.778$), caloric and cVEMPs findings ($p = 0.667$) and between cVEMPs and oVEMPs findings ($p = 0.583$).

Hypo activity to caloric stimulation in the affected ear is the most common finding in the present study, which is similar to the studies reported in the literature (Bergman & Stahle, 1967). In the present study, 7 out 9 (77.78%) affected ears showed hypo activity in caloric test. Hypo activity in caloric response is due to the damage to the hair cells in the horizontal semicircular canal (Murofushi et al., 2011). Only one subject had hyperactive responses for the caloric test. The hyperactive caloric responses in patients who suffer from Meniere's disease may be a transient phenomenon, caused by fluctuations of the vestibular condition, central compensation, age and/or mental state of the patients (Ikeda & Watanable, 1997).

Another significant finding in the present study was absence of both the cVEMPs as well as oVEMPs in 2 ears and absence of oVEMPs in one ear towards the contralateral side. Studies in the literature showed that second ear involvement in individuals with unilateral Meniere's disease was seen in 31% to 37% of cases (Thomas & Harrison, 1971; Green, Blum & Harner, 1991). Study by Lin et al. (2006) found that 27% of participants with unilateral Meniere's disease showed abnormal cVEMPs responses in the contralateral ear, which is similar to the present study. Histopathological studies of temporal bones of individuals with Meniere's disease showed that hydrops were more common in saccule and utricle compared to the semicircular canal (Okuno & Sando, 1987) so it can be concluded that, abnormal cVEMPs or oVEMPs responses may precede the symptoms in the contralateral ear, so VEMPs responses can be used to predict the chances of involvement of contralateral ear.

In the present study there was an association of the caloric and cVEMPs responses in 66.67 % of the cases, whereas there was an association of 55.56% between the cVEMPs and oVEMPs responses and 55.56% of the cases had association between the caloric and cVEMPs responses. However, the chi square test failed to show any statistically significant association between the two tests. The results on chi square test might have got affected because of the smaller sample size of the present study. Probably a larger data would have reflected any significant difference.

Significant association only between caloric and oVEMPs results, not between caloric and cVEMPs or between oVEMPs and cVEMPs have been reported in the literature (Murofushi et al. 2011). However, combining caloric test, oVEMPs and cVEMPs may provide localization of hydrops formation in the vestibular labyrinth (Huang, Wang & Young, 2011). For example, if normal caloric and abnormal oVEMPs findings indicate that, utricular macula is involved, while abnormal caloric and normal cVEMPs indicate that, semicircular canal is affected. Thus, one must carry out various vestibular tests in order to reach to a conclusive diagnosis in individuals with Meniere's disease.

Vestibular findings in Individuals with Vestibular Neuritis

Total 5 subjects diagnosed with vestibular neuritis participated in the study. All the subjects had unilateral vestibular neuritis. Table 7 shows the details of the subjects diagnosed with vestibular neuritis.

Vestibular evoked myogenic potentials

Cervical evoked myogenic potentials

Among the 5 participants with vestibular neuritis, cVEMPs recordings was absent in one ear (20%), and one ear (20%) had reduced amplitude with prolonged n23 latency. 3 ears (60%) showed normal cVEMPs responses. The latency and amplitude of cVEMPs responses for these 3 ears were similar to the individuals in control group. Kruskal-Wallis test was done to find out the difference between the p13 Latency, n23 latency and p13-n23 amplitude of the control group and vestibular neuritis and Kruskal-Wallis test revealed no significant difference between the latency and amplitude of the two groups ($p>0.05$).

Ocular - vestibular evoked myogenic potentials (oVEMPs)

Out of 5 subjects oVEMPs responses were absent in 2 ears (40%), and normal in 3 ears (60%) and the latency and amplitude measures in these 3 ears were not significantly different from the control group ($p>0.05$, Kruskal-Wallis test).

Electronystagmography

Central tests

Optokinetic test: All the participants had symmetrical optokinetic nystagmus in the optokinetic test.

Gaze test: All the participants had no nystagmus beats in the gaze test.

Peripheral tests

Positional test: Only one participant in the sub group B showed abnormal positional nystagmus in eyes closed condition for positions, namely, supine, supine with head turned to right, and supine with head turned to left.

Positioning test: One participant showed significant nystagmus beats in the positioning test.

Table 7: Clinical information of 5 participants with vestibular neuritis

				Air conduction thresholds (dBHL)						
S.No	Age (years)	Sex	Ear	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz	ABR
1	34	M	Right	5	10	10	15	5	10	N
2	38	F	Right	10	5	5	10	10	5	N
3	47	M	Left	10	10	5	15	10	10	N
4	42	M	Right	25	30	35	45	50	55	N
5	42	F	Right	5	10	5	15	10	10	N

N - Normal, M=Male, F= Female, ABR - Auditory Brainstem Responses.

Caloric test: Out of 5 ears, 4 ears showed hypo activity to caloric stimulation. Caloric responses in one ear were normal.

Association of Caloric test, cVEMPs and oVEMPs results in individuals with vestibular neuritis

In 60% (3 ears) of individuals, there was association of caloric test and oVEMPs findings. Two ears showed abnormal responses to both the tests. One ear showed normal findings in both the test. In 40% (2 ears) of individuals, there was dissociation of caloric and oVEMPss findings. One ear with normal oVEMPs responses showed abnormal caloric findings, one ear with normal caloric findings showed absent oVEMPs response.

In 60% (3 ears) of individuals, there was association of oVEMPs and cVEMPs findings. 2 ears showed normal responses to both the tests. One ear showed abnormal findings in both the test. In 40% (2 ears) of the individuals, there was disassociation between cVEMPs and oVEMPs responses. One ear with normal cVEMPs responses showed absent oVEMPs responses. One ear with absent cVEMPs responses showed normal oVEMPs responses.

In 20 % (1ear) of individuals, there was association of caloric test and cVEMPs findings. One ear showed abnormal responses to both the tests. In 80% (4 ears) of individuals, there was a disassociation between caloric test and cVEMPs responses. 3 ears with abnormal caloric test showed normal cVEMPs responses. One ear with normal caloric response showed abnormal findings in cVEMPs.

In the present study, 2 ears (40%) with vestibular neuritis showed abnormal findings in cVEMPs. Out of two eras with abnormal cVEMPs, one ear showed absent response and other ear showed reduced amplitude of p13-n23 complex and prolonged n23 latency indicating involvement of inferior vestibular nerve. Previous studies have reported 33.33% (Murofushi et al. 2011) and 39% (Murofushi, et al. 2001) abnormal cVEMPs findings in individuals with vestibular neuritis. In an another syudy in 134 patients with vestibular neuritis, Hong, Yeo, Kim and Cha (2008) found abnormal cVEMPs response in 36.6% of the participants, whereas Chiarovano et al. (2011) reported that 66.67% of participants showed abnormal cVEMPs responses in their study. The findings of the present study are almost similar to the study by Murofushi et al. (2001, 2011). The differences in different study might be due to involvement of different branches of vestibular nerve in individuals with vestibular neuritis (Aw, Fetter, Cremer, Karlberg & Halmagyi, 2001).

oVEMPs responses were absent in 2 ears (40%) indicating involvement of superior branch of vestibular nerve. Other studies have reported 75% (Chiarovano et al. 2011) and 100% (Murofushi et al., 2011) abnormal oVEMPs responses in individuals with vestibular neuritis. Variability in the results of different studies might be due to variable sample size and variable involvement of different branches of vestibular nerve in individuals with vestibular neuritis (Aw et al.2001). Combining the results of cVEMPs and oVEMPs, it can be concluded that the vestibular neuritis can have origin in inferior vestibular nerve alone or can have origin in superior vestibular nerve alone. The abnormal response of either cVEMPs or oVEMPs in individuals with vestibular neuritis is due to the involvement of the vestibular nerve. Individuals with vestibular neuritis have shown degenerative changes in the vestibular neuroepithelium, as well as in the vestibular nerve and the vestibular ganglion as shown in histopathological study of temporal bones in such individuals (Nodal, 1995), which might be a reason for absence of any recordable potentials from the vestibular nerves.

Prolonged latency was found in one participant and that might be due to the lesion at the nerve level which affects the conduction of impulses. Study by Hong et al. (2008) reported 59.2% of participants with prolonged p13 latency and 51% of participants with prolonged n23 latency. However, study by Murofushi et al. (2001) reported that none of the participants showed prolonged latency. The variability in results of different study due to, the variable extent of damage to the vestibular nerve in the participants of the different study.

None of the participants in the present study showed abnormal finding in the central test of ENG indicating involvement of only primary vestibular nerve in affected side, however abnormal central findings in ENG has been reported in the literature (Wennmo & Pyykko, 1982; Corvera & Davalos, 1985) due to the involvement of brainstem and cerebellum (Wennmo & Pyykko, 1982).

One of the participants in the present study had abnormal positional and positioning test findings. Wennmo and Pyykko (1982) reported 21 out of 30 individuals with vestibular neuritis having positional nystagmus, and reported that presence of positional nystagmus does not have any diagnostic significance in vestibular neuritis.

4 out of 5 ears (80%) showed, hypo activity caloric stimulation which is the most common findings reported in the vestibular neuritis (Wennmo & Pyykko, 1982). Hypo activity in caloric test indicates damage to the superior branch of vestibular nerve innervating horizontal semicircular canal (Wennmo & Pyykko, 1982).

Strong association of caloric test and oVEMPs findings have been reported in the literature (Murofushi et al. 2011). In the present study, 60% cases there was association between caloric and oVEMPs findings, 60%

cases had association between cVEMPs and oVEMPs and 20% cases had association between caloric test and cVEMPs finding. However, the chi square test failed to show any statistically significant association between the two tests. The results on chi square test might have got affected because of the smaller sample size of the present study. Probably a larger data would have reflected any significant association.

Vestibular neuritis can involve either superior branch or inferior branch of vestibular nerve or involvement of both the branches, based on involvement of different branch of vestibular nerve, vestibular neuritis can be superior or inferior or total vestibular neuritis (Murofushi et al. 2011).

Caloric test assess the functioning of superior branch of vestibular nerve innervating lateral semicircular canal, oVEMPs assesses majorly the functioning of superior branch innervating utricular macula and cVEMPs assess the functioning of inferior branch innervating saccular macula. Combining the 3 test will help us to determine involvement of different branches of vestibular nerve.

In the present study, one participant had abnormal findings in all 3 tests indicating total vestibular neuritis, one participant had abnormal findings in caloric and oVEMPs and normal cVEMPs responses indicating that inferior vestibular nerve is spared. Another participant had normal findings in caloric and oVEMPs, abnormal findings in cVEMPs responses indicating involvement of only inferior vestibular nerve. Remaining 2 participants had conflicting findings. They had abnormal caloric test results but normal oVEMPs and cVEMPs indicating involvement of superior nerve innervating lateral semicircular canal. Study by Kim et al. (2008) have reported that early recovery from abnormalities occur for otolith organs compared to the semicircular canals after vestibular neuritis. So the above 2 participants might have recovered the functioning of otolith organs with their respective branch of the vestibular nerve. However, such comment needs to be validated with studying recovery pattern of abnormalities of different tests after the onset of vestibular neuritis.

Vestibular findings in Individuals with idiopathic condition

Total 2 subjects (4 ears) in this group participated in this study. The diagnosis of these subjects could not be established by the Neurologist and the otolaryngologists.

Vestibular evoked myogenic potentials findings Individuals with Idiopathic condition

Cervical Vestibular evoked myogenic potentials

cVEMPs responses were abnormal in all 4 ears (100%). In one ear showed prolonged p13 and n23 latencies of cVEMPs. Other 3 ears showed absent cVEMPs responses.

Ocular - vestibular evoked myogenic potentials

Out of four ears, 3 ears showed absent oVEMPs responses (75%). oVEMPs responses was present in one ear and those responses were within normal limits.

Electronystagmography

Central tests

Optokinetic test: Both the participants had symmetrical optokinetic responses. No abnormality was noticed during the recording.

Gaze test:No nystagmus were recorded in right and left lateral gaze condition for both the participants. No abnormality was noticed during the recording.

Peripheral tests

Positional test: No nystagmus was present in positional test for both the participants

Positioning test: Both the participants had no nystagmus in the positioning test

Caloric test: In caloric testing one participant showed

Table 8: Clinical information of 2 participants with idiopathic condition

				Air conduction thresholds (dBHL)						
S.No	Age (years)	Sex	Ear	250 Hz	500 Hz	1kHz	2kHz	4kHz	8kHz	ABR
1	33	M	Right	15	20	15	25	20	25	N
2			Left	5	5	10	10	5	10	N
3	50	F	Right	5	10	5	5	10	15	N
4			Left	10	10	5	15	10	15	N

N - Normal, M - Male, F - Female, ABR - Auditory Brainstem Responses

Table 9: shows the relation between the symptoms exhibited by the individual and caloric, oVEMPs and cVEMPs findings

S. No	Symptoms	Number of subject with symptoms	Abnormal rate in Caloric test (%)	Abnormal and/or absent rate in oVEMPs (%)	Abnormal and/or absent rate in cVEMPs (%)
1	Lightheadedness or swimming sensation in the head	7	85.71	57.14	71.43
2	Blacking out or loss of consciousness	9	88.89	66.67	55.56
3	Tendency to fall.	13	92.31	92.31	61.54
4	Objects spinning or turning around you.	11	90.91	72.73	45.45
5	Sensation that you are turning or spinning inside.	4	100	100	100
6	Loss of balance when walking	13	92.31	76.92	61.54
7	Headache	5	60	60	40
8	Pressure in the head	5	80	60	60
9	Nausea or vomiting.	12	100	75	50

hypo activity to right ear and normal response in left ear stimulation. Another participant showed hyper activity response to right ear stimulations and response to left ear stimulation was normal.

First participant in the subgroup C had abnormal cVEMPs in both ears. Right ear had absent response and left ear had prolonged p13 and n23 latencies. oVEMPs responses was normal in right ear and absent in left ear. Caloric responses were normal in left ear and hypo activity in right ear. This participant had no otological diagnosis for the dizziness. This participant had no history or presence neurological or neuro muscular problems. From the above results it is clear that in this participant functioning of semicircular ocular reflex in right ear, left ear otolith-ocular reflex pathway and bilateral sacculo-collic pathways were affected.

Second participant had absent cVEMPs and oVEMPs responses bilaterally. Caloric test revealed normal response in the left ear and hyperactivity in the right ear. These participants also had no history or presence of any neurological or neuromuscular problems and have no otological diagnosis for dizziness. ENG, cVEMPs and oVEMPs revealed that in this participant functioning of semicircular ocular reflex in right ear, bilateral otolith-ocular reflex pathway and bilateral sacculo-collic pathways were affected.

In the above two participants, the diagnosis could not be established. The results from the above two participants clearly indicate the involvement, of multiple of anatomical structures. Because of the multiple anatomical structures involvement the sign

and symptoms exhibited by these clients may be very complex and hence the diagnosis could not be established. The combination of ENG, cVEMPs and two participants, it can be concluded that one particular test may not be able to find out the extent of lesion.

Association of symptoms exhibited by the individuals and test findings

Table 9 shows the relation between the symptoms exhibited by the individual and caloric, oVEMPs and cVEMPs findings. Only the section II of the dizziness questionnaire which contains main symptoms of individuals with dizziness was considered. In the present study, variable amount of association was obtained between the sign and symptoms exhibited by the clients and the different test results. A chi square test was administered to see any correlation between the sign and symptoms and the test results. Chi square test results revealed no significant correlation between any of the sign and symptoms with any of the test results ($p>0.05$), but for the symptom "Tendency to fall" was significantly correlated to the oVEMPs response ($p<0.05$). The lack of statistical association may be due to the small group of participants in the present study Based on these sign and symptoms exhibited by the client, and the test results obtained in the present study, one can narrow down the site lesion in the diagnosis of peripheral vestibular disorders. For example, the participants with nausea symptoms had 100% abnormal finding on caloric test and 75% abnormal finding on oVEMPs test, but in cVEMPs only 50% of the responses were absent. Since the caloric test and the oVEMPs assess the same superior vestibular nerves, it can be concluded with the above finding that lesions of superior vestibular nerves results in vomiting or nausea sensation in almost all the subjects. Similarly, subjects with subjective spinning sensations had 100% abnormal findings on the entire three test, based on which it can be concluded that the disorders of otolith organs or lateral canal results in spinning sensations. However, this will be very premature to generalize the results of the present study due to a smaller sample size.

Conclusions

Caloric test, cVEMPs and oVEMPs mainly assesses the functioning of semi circular canal ocular reflex pathway, sacculo-collic pathway and utriculo-ocular In vestibular dysfunction, one or more reflex pathways are affected. Since the above 3 tests assess the functioning of 3 different pathways, the combination of caloric test, cVEMPs and oVEMPs provides valuable information regarding localization of lesions in various peripheral vestibular disorders.

- ◇ This study provides information regarding the diagnostic significance of combination of ENG, cVEMPs, and oVEMPs in individuals with different vestibular dysfunction.
- ◇ This study provides basis for selection of the different kinds of vestibular rehabilitation in individuals with dizziness based on the affected reflex pathways.
- ◇ The study will help in identifying the exact site of lesion in subjects with different peripheral vestibular disorders.

References

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

Akin, F. W., & Murnane, O. D.Â (2001). Vestibular-evoked myogenic potentials: preliminary report. *Journal of the American Academy of Audiology, 12,* 445-452.

Akin, F. W., Murnane, O. D., & Medley, T. M.Â (2003). The effects of click and tone burst stimulus parameters on the vestibular evoked myogenic potential (VEMP).Â *Journal of the American Academy of Audiology, 14,* 500-508.

Akkuzu, G., Akkuzu, B., & Ozluoglu, L. N. (2006). Vestibular evoked myogenic potentials in benign paroxysmal positional vertigo and Meniere's disease. *European Archives of Otorhinolaryngology, 263,* 510-517.

Aw, S. T., Fetter, M., Cremer, P. D., Karlberg, M., & Halmagyi, G. M. (2001). Individual semicircular canal function in superior and inferior vestibular neuritis. *Neurology, 57,* 768-774.

Bakr, M. S., & Saleh, E. M. (2000). Electronystagmography: how helpful is it? *The Journal of Laryngology and Otology, 114*, 178-183.

Bergman, B., & Stahle, J. (1967). Caloric Reaction in Meniere's Disease. A Nystagmographic Study of 300 Patients.*Acta Oto-Laryngologica Supplements, 63,* 77-79.

Bhasali, S. A., & Honrubia, V. (1999). Current status of electronystagmography testing. *Otolaryngology-Head and Neck Surgery, 120*, 419-426.

Bohra, V., Sanju, H. K., & Sinha, S. K. (2012). A comparative study of cVEMP and oVEMP between dancers and non-dancers. Research paper presented at 44th ISHA conference in Hyderabad.

Boleas-Aguirre, M., Sanchez-Ferrandiz, N., Artieda, J., & Perez, N. (2007). Vestibular evoked myogenic potentials and benign paroxysmal positional vertigo. *Acta Otolaryngologica (Espanola), 58,* 173-177.

Bronstein, A. M., & Gresty, M. A. (1991). Compensatory eye movements in the presence of conflicting canal and otolith signals. *Experimental Brain Research, 85,* 697-700.

Chiarovano, E., Zamith, F., Vidal, P. P., & De Waele, C. (2011).Ocular and cervical VEMPs: a study of 74 patients suffering from peripheral vestibular disorders. *Clinical Neurophysiology, 122,* 1650-1659.

Coats, A. C. (1970). Central electronystagmographic abnormalities. *Archives of Otolaryngology, 92,* 43-53.

Colledge, N. R., Barr-Hamilton, R. M., Lewis, S. J., Sellar, R. J., & Wilson, J. A. (1997). Evaluation of investigations to diagnose the cause of dizziness in elderly people: a community based controlled study. *British Medical Journal, 313,* 788-789.

Corvera, J., & Davalos, R. L. (1985). Neurotologic evidence of central and peripheral involvement in patients with vestibular neuronitis. *Otolaryngology Head and Neck Surgery, 93,* 524-528.

De Waele, C., Tran Ba Huy, P., Diard, J. P., Freyss, G., & Vidal, P. P. (1999). Saccular dysfunction in Meniere's disease. *The American Journal Otology, 20,* 223-232.

Dizziness questionaure, Maryland hearing and Balance Center. Retrieved on September 2004. www.umm.edu/otolaryngology/dizziness_quest.doc

Green, J. D. Jr., Blum, D. J., & Harner, S. G. (1991). The longitudinal follow up of patients with Meniere's disease. *Otolaryngology Head and Neck Surgery, 104,* 783-788.

Halmagyi, G. M., & Curthoys, I. S. (1999). Clinical testing of otolith functions. *New York Academy of Sciences, 871*, 195-204.

Hong, S. M., Yeo, S., Kim, C., & Cha, T. (2008). The results of vestibular evoked myogenic potentials, with consideration of age-related changes, in vestibular neuritis, benign paroxysmal positional vertigo, and Meniere's disease. *Acta Oto-Laryngologica, 128,* 861-865.

Hulshof, J. H., & Baarsma, E. A. (1981).Vestibular investigations in Meniere's disease. *Acta Oto-Laryngologica, 92,* 75-81.

Huang, C. H., Wang, S. J., & Young, Y. H. (2011). Correlation between caloric and ocular vestibular evoked myogenic potential test results. *Acta Oto-Laryngologica, 132,* 160-166.

Ikeda, M., & Watanabe, I. (1997). Evaluation of hyperactive responses in inner ear diseases. *Journal for Oto-Rhino-Laryngologyand its Related Specialties, 59,* 326-331.

Jacobson, G. P., McCaslin, D. L., Piker, E. G., Gruenwald, J., Grantham, S. L., & Tegel, L. (2011). Patterns of Abnormality in cVEMP, oVEMP, and Caloric tests may provide topological information about vestibular impairment. *Journal of the American Academy of Audiology, 22,* 601-611.

Katayama, N., Yamamoto, M., Teranishi, M., Naganawa, S., Nakata, S., Sone, M., et al. (2010). Relationship between endolymphatic hydrops and vestibular evoked myogenic potential. *Acta Oto-Laryngologica, 130,* 917-923.

Kim, H. A., Hong, J. H., Lee, H., Yi, H. A., Lee, S. R., et al. (2008). Otolith dysfunction in vestibular neuritis: recovery pattern and a predictor of symptom recovery. *Neurology, 70,* 449-453.

Kirtane, M. V. (2009). Electronystagmography (ENG). In A. Biswas (Ed.), *Clinical audio-vestibulometry for otologists and neurologist* (4^{th} Edn) (pp.177-224). Mumbai: Bhalani publishing house.

Kumar, K. (2006). *Vestibular evoked myogenic potentials in normals and in individuals with dizziness.* Unpublished dissertation submitted to University of Mysore, Mysore.

Kumar, K., Sinha, S. K., Bharti, A. K., & Barman, A. (2006). *Comparison of Vestibular evoked myogenic potentials elicited by click & short duration one burst*. Research paper presented at 2^{nd} south zonal conference at on 14^{th} of May 2006, Kerala.

Lin, M. Y., Timmer, F. A., Oriel, B. S., Zhou, G., Guinan, J. J., Kujawa, S.G., et al. (2006). Vestibular evoked myogenic potentials can detect asymptomatic saccular hydrops. *Laryngoscope, 116,* 987-992.

Manzari, L., Burgess, A.M., & Curthoys, I.S. (2010). Dissociation between cVEMP and oVEMP responses: different vestibular origin of each VEMP? *Europian Archeives of Otorhinolaryngology, 267,* 1487-1489.

Murnane, O. D., Akin, F. W., Kelly, K. J., & Byrd, S. (2011). Effects of stimulus and recording parameters on the air conduction ocular evoked myogenic potential. *Journal of the American Academy of Audiology, 22,* 469-480.

Murofushi, T., Nakahara, H., Yoshimura, E., & Tsuda, Y. (2011). Association of air-conducted sound oVEMP findings with cVEMP and caloric test findings in patient with unilateral peripheral vestibular disorders. *Acta Oto-Layngologica, 131,* 945-950.

Murofushi, T., Shimizu, K., Takegoshi, H., & Cheng, P. W. (2001). Diagnostic value of prolonged latencies in the vestibular evoked myogenic potential. *Archives of Otolaryngology-Head & Neck Surgery, 127,* 1069-1072.

Nodal, J.B.Jr (1995). Vestibular Neuritis. *Otolaryngology Head and Neck Surgery, 112*(1), 162-72

Ochi, K., Ohashi, T., Nishino, H. (2001). Variance of vestibular-evoked myogenic potentials. *Laryngoscope, 111,* 522-527.

Okuno, T., & Sando, I (1987). Localization, frequency and severity of endolymphatic hydrops and the pathology of the labyrinthine membrane in Meniere's disease. *Annals of Otology, Rhinology and Laryngology, 96,* 438-445.

Ojala, M., Vaheri, E., & Juntenen, J. (1989). Electronystagmographic findings among 127 dizzy patients: correlation with the aetiology of dizziness. *Clinical Otolaryngology and Allied Sciences, 14,* 343-348.

Pflatz, C. T. (1984). Limits and possibilities of electronystagmography. *Laryngologie, Rhinologie, Otology, 63,* 511-516.

Park, H. J., Lee, I. S., Shin, J. E., Lee, Y. J., & Park, M. S. (2010). Frequency-tuning characteristics of cervical and ocular vestibular evoked myogenic potentials induced by air-conducted tone bursts. *Clinical Neurophysiology, 121*(1), 85-89.

Rauch, S. D., Zhou, G., Kujawa, S. G., Guinan, J. J., & Herrmann, B. S. (2004). Vestibular evoked myogenic potentials show altered tuning in patients with Meniere's disease. *Otology &Neurotology, 25,* 333-338.

Rosengren, S. M., Todd, N. P., & Colebatch, J. G. (2005). Vestibular-evoked extraocular potentials produced by stimulation with bone-conducted sound. *Clinical Neurophysiology, 116,* 1938-1948.

Schubert, M. C., & Shepard, N. T. (2008). Practical anatomy and physiology of the vestibulae system. In G. P. Jacobson & N. T. Shepard (Eds.), *Balance function assessment and management* (pp. 1-27). San Diego: Plural publishing inc.

Shi, M., Yu, X., Niu, H., & Lu, B. (1997).The value of electronystagmography in differential diagnosis of vertigo. *Bulletin of Hunan Medical University, 22,* 156-158.

Smulders, Y. E., Welgampola, M. S., Burgess. A. M., McGarvie, L. A., Halmagyi, G.M., & Curthoys, I. S. (2009).The n10 component of the ocular vestibular evoked myogenic potential (oVEMP) is distinct from the R1 component of the blink reflex. *Clinical Neurophysiology, 120*, 1567-1576.

Thomas, K., & Harrison, M. S. (1971). Long-term follow up of 610 cases of Meniere's disease. *Proceedings of the Royal Society of Medicines, 64,* 853-856.

Welgampola, M. S., Migliaccio, A. A., Myrie, O. A., Minor, L. B., & Carey, J. P. (2009). The human sound-evoked vestibulo- ocular reflex and its electromyographic correlate.*Clinical Neurophysiology, 120,* 158-166.

Wennmo, C., & Pyykko, I. (1982). Vestibular neuronitis: A Clinical and Electro-oculogruphic Analysis. *Acta Otolaryngologica, 94*, 507-515.

Young, Y. H., Huang, T. W., & Cheng, P. W.(2003). Assessing the stage of Meniere's disease using vestibular evoked myogenic potentials. *Archives of Otolaryngology-Head & Neck Surgery, 129,* 815-818.

Electrophysiological and Behavioral Assessment of Temporal Processing Abilities in Children with Dyslexia

[1]Satbir Singh & [2]Prawin Kumar

Abstract

Temporal processing refers to the time aspects of an auditory or acoustic signal. Temporal processing may be defined in several ways including determination of a sound source or " spatial percept," or determination of the pitch of a sound, and the perceptual segregation of two successive acoustic events. Temporal processing deficits have also been associated with learning disabilities. The present study investigates the performance of children with dyslexia and children without learning problem on different behavioral and electrophysiological tests of temporal auditory. The second aim to investigate the relationship between speech-evoked ABR responses and behavioral tests of temporal auditory processing in children with dyslexia and children without learning problem. Performance of 15 children with dyslexia (mean age 10.1 years) and 10 normal children without learning problem (mean age 9.5 years) were studied using behavioral tests of auditory temporal processing and speech-evoked ABR. The results indicate that all the children with dyslexia had deficits in brainstem timing. They also found to have poor performance on behavioral tests of temporal auditory processing. The present study concludes that BioMARK may be put to clinical use to identify the temporal processing deficits in difficult-to-test population and in monitoring the temporal processing abilities in children with dyslexia, following auditory training.

Keywords: *Dyslexia, speech-evoked ABR, temporal processing*

Introduction

(Central) Auditory Processing refers to the perceptual processing of auditory information in the central nervous system (CNS) and the neurobiological activity that underlies processing and gives rise to electrophysiological auditory potentials. It includes the auditory mechanisms that underlie the following abilities or skills: sound localization and lateralization; auditory discrimination; auditory pattern recognition; temporal aspects of audition, including temporal integration, temporal discrimination (e.g. temporal gap detection), temporal ordering, and temporal masking; auditory performance in competing acoustic signals (including dichotic listening); and auditory performance with degraded acoustic signals (ASHA, 1996; Bellis, 2003; Chermak & Musiek, 1997). (Central) Auditory Processing Disorder (C)APD refers to difficulties in the perceptual processing of auditory information in the CNS as demonstrated by poor performance in one or more of the above skills.

Temporal processing refers to the time aspects of an auditory or acoustic signal. Phillips (1995) defines temporal processing in several ways including determination of a sound source or " spatial percept," or determination of the pitch of a sound, and the perceptual segregation of two successive acoustic events. Temporal processing is important in the discrimination of duration and variations in pitch, which are critical to following the prosody of speech and music perception (Phillips, 1995).

Temporal processing deficits have also been associated with learning disabilities. Several authors have demonstrated that impaired temporal processing may result in language disorders, speech processing disorders and reading disorders (Merzenich, Jenkins, Johnston, Schreiner, Miller & Tallal, 1996; Tallal, Miller & Fitch, 1993). These investigators hypothesized that impaired temporal processing disrupts the normal development of an efficient phonological system and these phonological difficulties may result in language and reading disorders. Temporal processes are critical in a number of auditory functions including " auditory discrimination, binaural interaction, pattern recognition, localization/lateralization, monaural low-redundancy speech recognition, and binaural integration" (Show, Seikel, Chermak, & Berent, 2000). The underlying physiological neural mechanisms for temporal processing may be assessed by behavioral and electrophysiological means. Behavioral tests " stress" the auditory system by degrading the acoustic environment or signal by introducing background or speech noise or by filtering the signal. Behavioral tests may require multiple auditory processes such as attention, memory, and perception (Jirsa & Clontz, 1990).

Auditory-evoked potentials (AEPs) are commonly used to assess the temporal properties of the auditory system in a non-invasive fashion. Furthermore, AEPs have long been recognized as a reliable tool for providing objective information about the structural and functional integrity of the central auditory system (Hall, 1992; Kraus & McGee, 1992). Brainstem electrophysiological response elicited by speech stimuli may provide additional insight into the auditory processing abilities of some children with dyslexia. Speech-evoked ABR is a neurophysiologic response recorded to multiple presentations of a 40-ms synthetic /da/ syllable (Johnson,

[1]Email: satbirpgi2006@yahoo.com,
[2]Lecturer in Audiology, Email: prawin_audio@rediffmail.com

Nicol, & Kraus, 2005). The response manifests as a series of brief neural events that are time-locked to the onset, offset, and periodic information of the stimulus /da/. The response consists of two components: an onset response composed of Waves V, A, and C, and a sustained frequency-following response composed of Waves D, E, F, and O (Johnson et al., 2005).

Speech-evoked ABR has been used to investigate temporal processing deficits in children with language-based learning (including reading) disorders. Researchers have found that children with language-based learning disorders have abnormal speech-evoked brainstem responses (Banai, Nicol, Zecker, & Kraus, 2005; Banai et al., 2009). They concluded that the deficits observed may be due to a disruption at the brainstem level in timing and harmonic encoding.

Temporal processing may also be assessed behaviorally by tests of auditory temporal processing such as Masking Level Difference (MLD), Pitch Pattern Test, Duration Pattern Test, Random Auditory Gap Detection, and Time Compressed Speech Test. Behaviorally in the temporal processing ability of children at risk for APD will also be seen in results of electrophysiological tests (Musiek & Gollegly, 1988)

King, Lombardino, Crandell and Leonard (2003) investigated the performance of young adults with dyslexia on auditory processing tasks such as frequency pattern test (FPT) and duration pattern test (DPT) and found that 5 out of the 11subjects failed in both tests. Other studies have used different behavioral tasks such as same-different tasks (Tallal, 1980), identification of rapidly presented high-low frequency tones (Tallal, 1980; Farmer & Klein, 1993), or gap detection (Farmer & Klein, 1993) to investigate auditory processing in children and adults with reading disorders. They found significant difference in scores obtained by individuals with reading disorder and individuals without reading disorder. In contrast, Walker, Shinn, Cranford, Givens and Holbert (2002) found no significant differences in FPT scores between adults with dyslexia and a control group. These studies opine that children with learning disability have significant deficits in encoding of speech signal at brainstem as well as cortical level. They have poor performance on behavioral and electrophysiological tests of auditory processing.

There are various studies which reveal that a substantial proportion of children with auditory based learning problems such as dyslexia display abnormal encoding of speech signal as measured by the speech evoked auditory brainstem response (King, Warrier, Hayes, & Kraus, 2002; Warrier, Johnson, Hayes, Nicol, & Kraus, 2004; Banai et al., 2009;).

Furthermore, researchers have also investigated the performance of children with learning disability on behavioral tests of temporal auditory processing. (Farmer & Klein, 1993; King et al., 2003; Tallal, 1980).Hence, there is a need to check how the speech evoked auditory brainstem responses relate to the performance on behavioral tests of temporal auditory processing. Hence, the present study aimed to investigates the performance of children with dyslexia and children without learning problem on different behavioral tests of temporal auditory processing and speech-evoked ABR. It also aimed to investigate the relationship between speech-evoked ABR responses and behavioral tests of temporal auditory processing in children with dyslexia and children without learning problem.

Method

Participants

In the present study two groups was taken i.e. experimental group and control group. The experimental group consisted of 15 children with dyslexia (mean age 10.1 years) including 14 males and 1 female in the age range of 8 - 12 years. The diagnosis of dyslexia was made by an experienced speech and language pathologist/psychologist on the basis of following criteria. 1) scores below the normal range on the Early Reading Skills develop and standardized in Indian children by Loomba, (1995) and 2) Performance at least two grade levels lower than that expected of their chronological age on the scale, 'Appraisal of Kids with Specific Handicap in Arithmetic and Reading Activities' developed and standardized by Venkatesan (2002)

The control group consisted of 10 normal children without learning problem (mean age 9.5 years) including 7 males and 3 females in the age range of 8 - 12 years. All the participants in control group had good scholastic performance as per the detailed information gathered from the parents. All of them passed the screening checklist for auditory processing (SCAP) developed by Yathiraj and Mascarenhas (2004) indicating absent auditory processing disorder.

Those participants who had normal hearing sensitivity in the frequency range of 250 to 8000 Hz, normal click evoked ABR, normal middle ear function were included in both experimental and control group. The normal hearing was described as $\leq$ 15 dB HL for octaves frequencies from 250 to 8000 Hz and speech identification score (SIS) of $\geq$ 90% in both the ears. Participants who had peripheral hearing loss, clinically abnormal/absent click-evoked ABR, any middle ear pathology, limited intellectual capacity, and attention deficit hyperactivity disorder (ADHD) were excluded from the experimental as well as from the control group.

Materials and procedure

The testing was carried out in two phases. The first phase was a preliminary screening session which included auditory electrophysiological testing. The second phase consisted of a behavioral central auditory testing. All the behavioral as well as electrophysiological tests were carried out in the sound treated room where the noise level was as per the guidelines in ANSI S3.1 (1991).Informed consent was obtained from all participants and their parents.

The phase 1 consisted of a complete hearing screening, click-evoked ABR testing, and speech-evoked ABR testing. Screening Checklist for Auditory processing (SCAP) developed by Yathiraj and Mascarenhas (2004) was administered which consists of twelve questions having the symptoms of deficits in auditory processing (Auditory perceptual processing, Auditory memory and others). The scoring was done as 'Yes' or' No'. Each answer was marked "yes" carried one point and "no" carried zero point. Those children who scored less than 50% (< 6/12) in SCAP were considered for the study. Pure-tone thresholds were obtained using calibrated double channel clinical audiometer (Orbiter-922) with TDH-39 headphones at octave frequencies between 250 Hz to 8000 Hz for air conduction and between 250 Hz to 4000 Hz for bone conduction through modified Hughson Westlake procedure (Carhart & Jerger, 1959). Immittance audiometry was carried out using a calibrated middle ear analyzer (GSI-tympstar) with a probe tone frequency of 226 Hz. Ipsilateral and contralateral acoustic reflexes thresholds were measured for 500, 1000, 2000, and 4000 Hz.

Electrophysiological testing which includes click and speech-evoked ABR was carried out using Biologic Navigator Pro EP system (version 7.0). Click evoked ABR testing was performed to verify normal transmission of auditory stimuli through the brainstem auditory pathway. For recording the click evoked and speech evoked ABR the clients were asked to sit on a reclining chair. The site of electrode placement was cleaned thoroughly with skin abrasive to reduce the skin-electrode impedance to less than 5 kΩ. Electrodes were placed with the help of skin conduction paste at Cz (non-inverting), with inverting electrode at the mastoid of test ear (M1) and the ground electrode at contralateral mastoid (M2). Responses were obtained for both ears to rarefaction click stimuli presented at 90 dB nHL with an online filter of 100 - 3000 Hz. Two thousand stimulus repetitions were collected at a rate of 11.1/sec. All participants in control as well as in experimental group had click-evoked ABR within the clinical norms.

Once the normal click-evoked ABR and normal hearing sensitivity was conformed, all participants underwent speech-evoked ABR testing. Electrodes (impedance < 5 kΩ) were placed with the help of skin conduction paste at Cz (non-inverting), with inverting electrode at the mastoid of test ear (M1) and the ground electrode at contralateral mastoid (M2). Responses were obtained for both ears to 40-ms speech like /da/ stimulus with alternating polarity. Stimuli were presented at 90 dB SPL with an online filter of 100 - 2000 Hz. Two thousand stimulus repetitions were collected at a rate of 10.9/sec.

Speech-evoked ABR composed of the transient and the sustained responses (also known as frequency following responses). Transient responses consists of peak V and A, whereas the sustained responses consist of peak D, E, F and O. In the present study both transient as well as sustained responses were evaluated. Two repeatable recordings were obtained in order to verify response replicability. Peaks were marked on the resultant waveform which was obtained after 'weighted-add' of two replicable waveform. In order to get spectral components of speech-evoked ABR, waveforms were first converted into "ASCII" formant using the software called 'AEP to ASCII' (version 1.6.0). ASCII formant data was then analyzed using the MATLAB platform and software 'BRAINSTEM TOOLBOX (Skoe & Kraus, 2010).

In phase 2, a test battery of three behavioral tests used in the diagnosis of temporal processing abilities in children with (C) APD was used. It includes Duration Pattern Test (Gauri, 2003), Pitch Pattern Sequence Test (Shivani, 2003) and Gap Detection Test (Shivaprakash, 2003). Test stimuli was routed from a personal computer (PC) with Intel Celeron processor through a two-channel clinical audiometer (Madsen OB-922) with TDH-39 headphones at an intensity level of 40 dB SL (re: PTA) binaurally. Initially 1 kHz calibration tone was presented to the subject's ear through TDH-39 earphone and V-U meter was adjusted to show "0" reading. Practice items were presented before the beginning of each test to ensure understanding of the task. Order of test administration within the behavioral central auditory test battery was counter balanced across participants to control for possible order effects. Listening breaks were given periodically throughout the testing session or as per the participant's request. The duration of the test session was approximately 60 to 90 minutes.

Statistical Analyses

Scores of behavioral as well as electrophysiological tests obtained from 10 normal children without learning problem and 15 children with dyslexia were analyzed using SPSS (version 19) software. Beside descriptive statistics, parametric test Multivariate Analysis of Variance (MANOVA) was used to compare the performance between the two groups. Further, in order to find the correlation between electrophysiological and behavioral tests, Karl Pearson correlation was used.

Results

The data collected from dyslexic children and children without learning problem were tabulated. Both descriptive and inferential statistical analysis was carried out

Table 1: Mean and SD values of amplitude of sustained responses of speech-evoked ABR of experimental and control group

Parameters	Control group (N = 20)		Experimental group (N = 30)		p Value
	Mean	S.D	Mean	S.D	
F0 (μV)	8.51	5.32	8.21	4.67	0.836
F1 (μV)	1.81	0.74	0.79	0.45	0.000***
F2 (μV)	0.58	0.08	0.02	0.15	0.000***

Note:p< 0.05; p < 0.01**; p < 0.001***; N = number of ears*

Table 2: Mean and standard deviation (SD) values of speech-evoked ABR parameters of control and experimental group

Parameters	Control group (N = 20)		Experimental group (N = 30)		p Value
	Mean	S.D	Mean	S.D	
Wave V (ms)	6.87	0.30	8.25	1.64	0.001**
Wave A (ms)	7.94	0.52	9.70	2.65	0.005**
Wave C (ms)	17.75	0.73	20.45	3.28	0.001**
Wave D (ms)	23.09	0.84	25.98	4.43	0.006**
Wave E (ms)	31.61	0.73	35.09	4.32	0.001**
Wave F (ms)	39.67	0.71	43.36	3.80	0.000***
Wave O (ms)	48.64	0.80	50.75	3.20	0.006**
V/A slope (μV/ms)	0.01	0.12	-0.03	0.05	0.592

Note: p< 0.05; p < 0.01**; p < 0.001*** ;N = number of ears*

for speech-evoked ABR as well as for the auditory temporal processing tests.

Results of Speech-evoked ABR

The speech-evoked ABR data was analyzed in terms of latency and amplitude. The waves V, A, C, D, E, F and O of speech-evoked ABR were identified and their latencies and amplitude were noted. Fast-Fourier Transform (FFT) was carried out to find the amplitude of F0, F1 and higher harmonics (F2) frequency components elicited by syllable /da/ of 40 msec.

Speech-evoked ABR could be recorded from all dyslexic children and children without learning problem. It was observed that the overall waveform morphology was poorer and increased (prolonged) in latency in dyslexic children (Figure 2) as compared to children without learning problem (Figure 1)

Visual inspection of individual data revealed that out of 30 ears wave C was absent in 3 ears (20%) and wave O was absent in 2 ear (13.3%) of the dyslexic children. However in children without learning problem all the waves were present. Results of MANOVA (Table 1) revealed that latencies of waves - were significantly prolonged in dyslexic children as compared to children without learning problem ($p < 0.05$). These findings suggest that dyslexicchildren showed abnormal encoding of speech signal at the brainstem level.

Results of MANOVA (Table 2) showed a significant reduction in amplitude of harmonics (F1 & F2) in dyslexic children as compared to the children without learning problem ($p < 0.001$). However, there was no significant difference in the amplitude of fundamental frequency (F0) between dyslexic children and children without

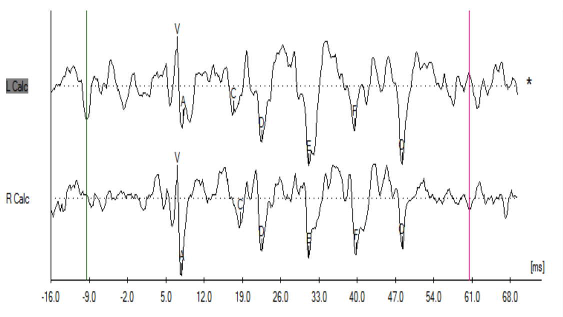


Figure 1: Sample waveform of speech-evoked ABR in children without learning problem.

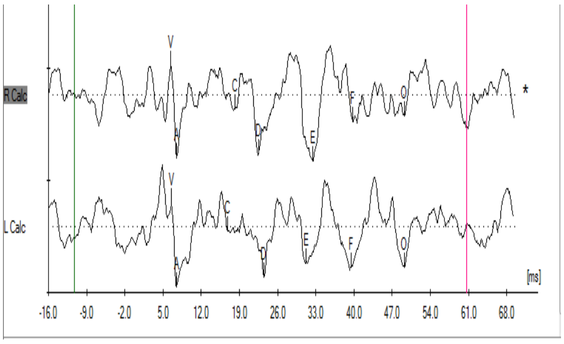


Figure 2: Sample waveform of speech-evoked ABR in dyslexic children.

learning problem ($p > 0.05$). These findings suggested that the spectral information within the F0 region of the response remain robust in the dyslexic children.

Results of Auditory Temporal Processing tests

For behavioral tests descriptive statistics which included means and standard deviation was carried out for both experimental and control groups. A parametric test, MANOVA was used to check if there were significant differences between means of the two groups.

Table 3 showed the mean scores of behavioral tests of temporal processing in children without learning problem and in dyslexic children. The scores of GDT, PPST, and DPT were significantly reduced in dyslexic children as compared to children without learning problem ($p < 0.05$). In dyslexic children, the mean gap detection threshold was 4.40 msec, whereas in children without learning problem, the mean gap detection threshold was found to be 11.23 msec, which was significantly higher in dyslexic children as compared to children without learning problem. This depicts the reduced temporal resolution ability in children with dyslexia.

Relationship between Speech-Evoked ABR and Behavioral Tests of Temporal Auditory Processing

Karl pearsons correlation was used to check if there were any correlation between speech-evoked ABR (wave V, V/A slope and spectral components of sustained portion i.e. F0, F1 and F2) and behavioral tests of temporal auditory processing- GDT, PPST, DPT in both dyslexic children as well as in children without learning problem.

Results of Karl pearsons correlation (Table 4) showed a significant negative correlation between transient response (wave V) of speech-evoked ABR and PPST scores in dyslexic children ($r = -0.587$, $p < 0.05$). This suggests that latency of wave V of speech-evoked ABR tend to be prolonging (abnormal) when there is decrease (poor) in PPST scores and vice-versa. Wave V was also found to be have a non-significant negative correlation with GDT scores ($r = -0.25$, $p >$

Table 3: Mean and SD values of behavioral tests scores of experimental and control group

	Control group (N = 10)		Experimental group (N = 15)		p Value
Parameters	Mean	S.D	Mean	S.D	
GDT	4.40	1.34	11.23	1.63	0.000***
PPST	23.40	1.17	16.26	3.10	0.000***
DPT	23.30	2.21	15.33	2.76	0.000***

Note: p< 0.05; p < 0.01**; p < 0.001***; N = number of ears*

Table 4: Correlation values of speech-evoked ABR's parameters (both transient & sustained responses) and behavioral tests scores of experimental group

Parameters	GDT		PPST		DPT	
	r value	p value	r value	p value	r value	p value
Wave V	- 0.25	0.37	- 0.08	0.02*	0.27	0.33
V/A slope	0.24	0.38	0.03	0.91	- 0.08	0.77
F0	- 0.29	0.28	- 0.53	0.07	0.39	0.14
F1	- 0.02	0.41	0.38	0.15	0.39	0.14
F2	- 0.27	0.33	0.39	0.14	0.55	0.65

Note:p< 0.05; p < 0.01**; p < 0.001***; r = correlation coefficient*

0.05) and non-significant positive correlation with DPT scores (r = 0.27, $p > 0.05$). Moreover, a well negative but non-significant negative correlation was also found between GDT and all transient and sustained responses of speech-evoked ABR except V/A slope, where the correlation was positive and non-significant ($p > 0.05$). However, in children without learning problem, no correlation was found between speech-evoked ABR and behavioral tests of temporal auditory processing.

Discussion

Latency and amplitude of Speech-evoked ABR in children with dyslexia

The results of present study are consistent with the previously published studies (Abrams, Nicol, Zecker & Kraus, 2006; Banai et al., 2005; Banai et al., 2009; Johnson et al., 2005; King et al., 2002; Russo et al., 2004; Wible , Nicol & Kraus, 2005), which have shown that children with learning disorders, such as dyslexia, have been found to exhibit delayed peak latencies for waves V, A, C, and O and a shallow slope for V/A slope indicating abnormal brainstem timing to speech signal.

However, in the present study no statistically significant difference was found in V/A slope between dyslexic children and children without learning problem ($p > 0.05$). This could be due to heterogeneous nature of dyslexia and small sample of dyslexic population taken in the present study. Billiet and Bellis (2011) found that children with normal brainstem timing who met the diagnostic criteria for (C)APD using behavioral measures did exhibit some abnormalities in the temporal, rather than spectral, elements for their speech-evoked ABR responses, although overall speech-evoked ABR scores were well within normal range.

Other studies (Banai et al., 2005; Banai et al., 2009) have shown that children with dyslexia have significantly reduced spectral information for speech-evoked ABR responses. It has also been found that not all children with dyslexia show deficits in temporal processing. There are some studies which failed to demonstrate any auditory temporal processing deficit in children with dyslexia (Bretherton & Holmes, 2003; Brier, Fletcher, Foorman, Klaas & Gray, 2003; Mody, Studdert-Kennedy & Brady, 1997; Schulte-Korne, Deimel, Bartling, & Remschmidt, 1999; Watson & Miller, 1993; Watson & Kidd, 2002; Ziegler, Pech-Georgel, George, & Lorenzi, 2009).

In the present study it has also been found that amplitude of harmonics (F1 & F2) was reduced in dyslexic with changes amplitude of fundamental frequency (F0) children. These findings are consistent with the findings of (Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001, Wible, Nicol & Kraus, 2004). They found that in the presence of noise or rapid stimulation, spectral cues present in F1 remain robust but it diminished in F0 region in the children with learning disability. Thus it can be concluded that children with dyslexia demonstrate selective disruptions in brainstem encoding of the F1 and F2 characteristics of the speech signal, whereas F0 information remain relatively intact. Banai et al. (2009) reported that these deficits may be due to a disruption at the brainstem level in timing and harmonic encoding

Temporal auditory processing tests in children with dyslexia

Results of behavioral tests are consistent with the previously published studies (Baldeweg, Richardson, Watkins, Foale & Gruzilier, 1999; Dougherty, Cynader, Bjornson, Edgell & Giaschi, 1998; Schulte-KĂśrne et al., 1998). Results of these studies have shown that children with dyslexia have significant deficits in frequency discrimination task and the tone-in-noise detection task resulting in poor temporal resolution.

Ingelghem et al. (2001) also assessed temporal processing in individuals with dyslexia by means of two psychophysical threshold tests - Gap detection in broad band noise and Frequency Modulation (FM) detection. They concluded that the results of the temporal processing assessment were statistically poorer in dyslexic children as compared without learning problem. A possible neurophysiologic explanation for this observed deficit in auditory temporal resolution is that dyslexic readers have a prolonged refractive period in their neurological firing pattern. This may be the result of a slower transmission time of neural information (Stein & Walsh, 1997).

In the present study, we also found that dyslexic children had poor performance on temporal patterning/ordering tasks. These findings are consistent with the study done by Tallal (1980), in which it was found that children with dyslexia had poor performance on temporal tests (Sequencing Test, Rapid Perception test, and Same-Different Discrimination Test) performed by the group with reading and writing disorder as compared to children without reading and writing problem.

Hari and Keisila (1996) studied the temporal processing in dyslexic children with trains of binaural clicks which led to illusory movements at short click intervals. They found that children in control group, the illusion disappeared at intervals exceeding 90-120 ms, while in dyslexics it persisted up to intervals of 250-500 msec. They concluded that dyslexic children seem to have deficit in the temporal processing of rapid sequences. In the another study, Murphy and Schochat (2009) also studied auditory temporal processing in Brazilian children with dyslexia and reported that the group of children with dyslexia showed poor performance on the temporal auditory processing tests developed by Tallal and Piercy (1973).

The neural basis of the timing deficit is unclear. According to researchers (Llinas, 1993; Merzenich et al., 1996) one possibility is that the rate necessary for filling in and reading out the sensory buffers, is slower in dyslexics than in normal subjects. Such slowing is thought to be associated with decreased frequency and synchrony of intrinsic neuronal oscillations, both in the cortex and in the thalamocortical system.

Relationship between speech evoked ABR and different temporal processing tests

The correlation findings between speech-evoked ABR and behavioral tests score indicate that abnormal brainstem responses to speech stimuli may also be an indicator of poor responses at the cortical level. This can be explained by the study done by Nicol, and Kraus (2004). They found that broader V/A slopes were correlated with an increased vulnerability of the cortical response to the effects of background noise. In the another study, Banai et al. (2005) also found that children with language-based learning disorders who exhibited abnormal speech-evoked brainstem responses also had reduced speech-evoked mismatch negativity responses compared with children without learning problem.

Conclusions

To conclude, in the present study it has been found that the BioMARK was able to detect brainstem timing deficits in children with dyslexia. Hence BioMARK may be put to clinical use to identify the temporal processing deficits in difficult-to-test population and in monitoring the temporal processing abilities in children with dyslexia, following auditory training. Moreover, it has also been found that there is a relationship between brainstem timing and cortical processing as depicted by correlation findings between speech-evoked ABR and behavioral tests of temporal auditory processing. However, this relationship may not be present in all children with dyslexia and may not be seen for all tests of auditory processing as depicted by the findings of present study. Moreover, central auditory processing is not affected in all children with reading deficits and reading deficits are not exhibited by all children with (C)APD in comparison to children without learning problem. This demonstrates the heterogeneous nature of this disorder.

References

Abrams, D. A., Nicol, T., Zecker, S. G., & Kraus, N. (2006). Auditory brainstem timing predicts cerebral asymmetry for speech. *The Journal of Neuroscience, 26*, 11131-11137.

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

American Speech-Language-Hearing Association. (1996). Central auditory processing: Current status of research and implications for clinical practice. *American Journal of Audiology, 5*, 41-54.

Baldeweg, T., Richardson, A., Watkins, S., Foale, C., & Gruzilier, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Annals of Neurology, 45*, 495-503.

Banai, K., Nicol, T., Zecker, S. G., & Kraus, N. (2005). Brainstem timing: Implications for cortical processing and literacy. *Journal of Neuroscience, 25*, 9850-9857.

Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex, 19*, 2699-2707.

Bellis, T. J. (2003). *Assessment and management of central auditory processing disordersin the educational setting: From science to practice* (2nd ed.). Clifton Park, NY: Delmar Learning.

Bretherton, L., & Holmes, V. M. (2003).The relationship between auditory temporal processing, phonemic awareness, and reading disability. *Journal of Experimental Child Psychology, 84*, 218-243.

Brier, J. I., Fletcher, J. M., Foorman, B. R., Klaas, P., & Gray, L. C. (2003). Auditory temporal processing in children with specific reading disability with and without attention deficit/hyperactivity disorder. *Journal of Speech Language and Hearing Research,46*, 31-42.

Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders, 24*, 350-345.

Chermak, G. D., & Musiek, F. E. (1997).*Central auditory processing disorders: New perspectives*. San Diego, CA: Singular.

Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology, 112*, 758-767.

Dougherty, R. F., Cynader, M. S., Bjornson, B. H., Edgell, D., & Giaschi, D. E. (1998). Dichotic pitch: A new stimulus distinguishes normal and dyslexic auditory function. *Neuroreport, 9,* 3001-3005.

Farmer M. E., & Klein, R. (1993).Auditory and visual temporal processing in dyslexic and normal readers. *Annals of New York Academy of*

Science, *682*, 339-341.

Gelfand, S. (1998).*Hearing:An introduction to psychophysical and physiological acoustics* (3rd Ed.). New York: Marcel Dekker.

Gauri, D.T. (2003). *Development of Norms on Duration Pattern Test*. Unpublished Independent Project, University of Mysore, Mysore.

Hall, J. W. (1992). *Handbook of auditory evoked responses*. Needham Heights, Mass: Allyn & Bacon.

Hari, R, H., & Keisila, P. (1996). Deficits of temporal auditory processing in dyslexic adults. *Neuroscience Letters,205*, 138-140.

Ingelghem, M., Van Wieringen, A., Wouters, J., Vandenbussche, E., Onghena, P., & Ghesquiere, P. (2001). Psychophysical evidence for a general temporal processing deficit in children with dyslexia. *Neuroreport, 12*, 3603-3607.

Jirsa, R. E, & Clontz, B. C. (1990). Long latency auditory event related potentials from children with auditory processing disorders. *Ear and Hearing, 11*, 222-232.

Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brainstem response to speech: A biological marker of auditory processing. *Ear and Hearing, 26*, 424-434.

Kraus, N. & McGee, T. (1992).Electrophysiology of the human auditory system. In A. Popper & R. Fay (Eds.) *The mammalian auditory system.* New-York: Springer Verlag, 335-404.

King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002).Deficits in auditory brainstem encoding of speech sounds in children with learning problems.*Neuroscience Letters, 319*, 111-115.

King W. M., Lombardino, L. J., Crandell, C. C., & Leonard, C. M. (2003). Co-morbid auditory processing disorder in developmental dyslexia. *Ear and Hearing*, *24*, 1448-1456.

Llinas, R (1993). Is dyslexia a dyschronia ?*Annals of the New York Academy of Science, 682*, 48-56.

Loomba, K. (1995). *Descriptive analysis of the sequential progression of English reading skills among Indian children*. An Unpublished Master's dissertation, University of Mysore, Mysore.

Menon, D. (2007). *Speech-evoked Mismatch Negativity: An objective measure of auditory perceptual deficits. An Unpublished doctorate thesis,* University of Mysore, Mysore.

Merzenich, M. M., Jenkins, J. W., Johnston, P., Schreiner, C., Miller, S. L., &Tallal, P. (1996). Temporal processing deficits in language learning impaired children ameliorated by training. *Science*, *271*, 77-81.

Mody, M., Studert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology, 64*, 199-231.

Murphy, C. F. B., & Schochat, E. (2009). How auditory temporal processing deficits relate to dyslexia. *Brazilian Journal of Medical and Biological Research*, *42*, 647-654.

Musiek, F. E., & Gollegly, K. (1988). *Maturational considerations in the neuroauditory evaluation of children*. In. F. Bess (Ed.), Hearing impairment in children (pp. 231-252). Parkton, MD: York Press.

Nicol, T., & Kraus, N. (2004). Speech-sound encoding: physiological manifestations and behavioral ramifications. *Clinical Neurophysiology Supplement, 57*, 624-630.

Phillips, D. (1995). Central auditory processing: A view from auditory neuroscience. *The American Journal of Otology, 16*, 338-352.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115* , 2021-2030.

Schulte-Korne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1999). Pre-attentive processing of auditory patterns in dyslexic human subjects. *Neuroscience Letters, 276,* 41-44.

Shivani, T. (2003). *Maturational effect of pitch pattern Sequence Test.* An Unpublished Independent Project, University of Mysore, Mysore.

Shivaprakash, S. (2003). Gap Detection Test-Development of Norms. An Unpublished Independent Project, University of Mysore, Mysore.

Show, R. L, Seikel, Chermak, G. D., & Berent, M. (2000). Central Auditory Processes and Test Measures: ASHA 1996 Revisited. *American Journal of Audiology*: *A Journal of Clinical Practice, 9*, 63-68.

Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: a tutorial. *Ear and Hearing, 31*, 302-324.

Stein, J., & Walsh, V. (1997).To see but not to read; the magnocellular theory of dyslexia. *Trends in Neuroscience, 20*, 147-152.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*, 182-198.

Tallal, P., Miller, S., & Fitch, R. H. (1993). *Neurobiological basis of speech: A case for the preeminence of temporal processing*. In P. Tallal, R. R. Galaburda, R. R. Llins & C. von Euler (Eds.), Temporal information processing in the nervous system (pp. 27-47). New York: The New York Academy of Sciences, DI. 682.

Venkatesan, S. (2002). '*Appraisal of Kids with Specific Handicap in Arithmetic and Reading Ac-*

tivities. An Unpublished Project, University of Mysore, Mysore.

Walker, M., Shinn, J., Cranford, J. L., Givens, G., & Holbert, D. (2002). Auditory temporal processing performance of young adults with reading disorders. *Journal of Speech, Language and Hearing Research*, *45*, 598-605.

Watson, C. A., & Kidd, G. R. (2002).On the lack of association between basic auditory abilities, speech processing, and other cognitive skills. *Seminars in Hearing, 23*, 83-93.

Watson, B. U., & Miller, T. K. (1993).Auditory perception, phonological processing, and reading ability/disability. *Journal of Speech and Hearing Research*, *36*, 850-863.

Wible, B., Nicol, T., & Kraus, N. (2004). Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychology, 67*, 299-317.

Wible, B., Nicol, T., & Kraus, N. (2005).Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain, 128*, 417-423.

Yathiraj, A., & Mascarenhas, K. (2004).Audiological profile of children with suspected auditory processing disorder. *Journal of Indian Speech and Hearing Association, 19*, 5-14.65

Ziegler, J., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech perception in noise deficits in dyslexia. *Developmental Science, 12*, 732-745.

Effect of Spectral Bandwidth and Spectral Integration on Speech Perception in Listeners with Normal Hearing, Cochlear Hearing Loss and Auditory Dys-Synchrony

[1]Seby Maria Manuel & [2]Animesh Barman

Abstract

The present study aimed at finding whether the type of hearing loss has any effect on the bandwidth required to achieve minimum amount of speech identification scores with low and high center frequencies and also ability to integrate information from these two bandwidths. Participants included 29 individuals with normal hearing, 12 individuals with cochlear hearing loss and 17 individuals with auditory dys-synchrony. Two CSBs (Criterion Speech bandwidth) were obtained for all the participants. The first one for the low center frequency of 500 Hz and the second one for the higher center frequency of 2500 Hz. To determine the spectral integration abilities, words having both the CSBs were presented to the participants. Results showed that individuals with auditory dys-synchrony failed to achieve criterion score even at maximum bandwidth at 500 Hz center frequency. All the three groups differed significantly from each other for the normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 2500 Hz. Individuals with auditory dys-synchrony showed significantly lower spectral integration scores from the other two groups and they showed two types of spectral integration scores. One group showed reduced spectral integration scores compared to the criterion score obtained at 2500 Hz center frequency. Another group showed marginal improvement. It was seen that individuals with cochlear hearing loss were as good as normal hearing individuals in their ability to combine the information across different frequency bands.

Keywords: *Spectral integration, Criterion Speech Bandwidth, center frequency*

Introduction

Speech is a complex signal. The components of speech vary in terms of frequency and intensity over time. Approximately 95% of the frequency components in speech lie between 300 Hz and 3000 Hz (Hamernik & Davis, 1988). To perceive and understand speech one needs to have normal hearing sensitivity within this frequency range. Hearing loss at any frequency within this frequency range will affect speech perception. The impact of hearing loss on speech perception is based on both type and configuration of hearing loss.

It is well established that different speech sounds have predominantly different energies across frequencies. For example nasals are predominantly lower in frequency, whereas fricatives are more of high frequency speech sounds. Thus, individual with low frequency hearing loss will have difficulty perceiving nasals and individuals with high frequency hearing loss will not be able to get the important features which are necessary to perceive fricatives. Hence, all these individuals would fail to comprehend speech.

Similarly, type of hearing loss has also different perceptual consequences (Zeng & Liu, 2006). A conductive type of hearing loss which is thought to attenuate the acoustic signal reaching to the cochlea is likely to have less impact on speech perception, whereas cochlear hearing loss would show deterioration in speech perception with the increase in severity of hearing loss. This is probably due to the loss of OHC's in the cochlea, which is responsible for fine discrimination. As the fundamental frequency, formant frequency and frequency transition are important features to understand speech, perception of these features will be affected due to lack of sharper tuning as a result of OHC damage. People with cochlear hearing loss usually have auditory filters that are broader than normal (Glasberg & Moore, 1986; Tyler, 1986). This means that their ability to determine the spectral shape of speech sounds and to separate components of speech from background noise is reduced. Impaired frequency resolution has been identified as the main reason for speech perception deficits in cochlear hearing loss with greater than moderate degree of hearing loss (Glasberg & Moore, 1989).

Another reason for impaired speech perception can be, reduced phase locking in these individuals. This may be due to fact that the propagation time of the travelling wave along the basilar membrane can be affected by the cochlear damage and this may disrupt the processing of temporal information by central mechanisms (Leob & White, 1983).

People with cochlear hearing impairment often complain that their greatest problem is understanding speech when background noise is present. The hearing impaired needs a higher signal-to noise ratio (SNR) to achieve the same level of performance (Plomp, 1994). This increase in signal to noise ratio ranged from 2.5 dB for mild hearing loss to 7 dB for moderate to severe

[1]Email: sebymanuel88@gmail.com,
[2]Reader in Audiology, Email: nishiprerna@yahoo.com

hearing loss. An even larger SNR is required when the noise is fluctuating (Plomp, 1994).

Auditory dys-synchrony is another hearing disorder that has unique pathologies and perceptual consequences (Starr, Picton, Sininger, Hood & Berlin, 1996). It is a disorder characterized by abnormal or absent auditory brainstem responses (ABRs) and the presence of otoacoustic emissions (OAEs) and/or cochlear microphonics (CMs), indicating normal functioning of the outer hair cells (OHCs) (Starr et al., 1996).

It is difficult to localize the exact cause for auditory dys-synchrony. There may be multiple underlying causes (Rance, 2005). Auditory dys-synchrony (AN) may result from a loss of inner hair cells (IHC), dysfunction of the IHC-nerve synapses, neural demyelination, axonal loss or a possible combination of multiple sites. These pathologies may be present with the traditional cochlear loss involving outer hair cells and/or central processing disorders involving the brainstem and cortex, complicating the classification of auditory dys-synchrony (Rapin & Gravel, 2003).

One major characteristic of AN is an impaired capacity for temporal processing and difficulty in speech understanding, particularly in noise, that is disproportionate to the degree of hearing loss measured by pure-tone thresholds (Rance, Cone- Wesson, Wunderlich & Dowell, 2002; Rance, McKay & Grayden, 2004, Zeng, Kong, Michalewski & Starr, 2005). Zeng and Liu (2006) said that these individuals have poor pitch processing at low frequencies, excessive masking in noise, and inability to process interaural timing information. Most of the individuals with auditory dys-synchrony have a raising pattern of hearing loss indicating a low frequency hearing loss. This is mainly due to the auditory nerve fibers which are getting affected in them since the low frequency fibers are the longest ones they have more chances of getting involved and this results in poor pitch processing at low frequencies. Zeng, Oba, Garde, Sininger and Starr (1999) studied the frequency discrimination abilities of these individuals across frequencies and found that they have very poor discrimination at low frequencies. Even at signal to noise ratios of 10 to 15 dB, individuals with auditory dys-synchrony found it difficult to perceive speech which is due to the excessive masking.

Several studies have tried to explain the reasons for poor speech perception abilities, especially in the presence of noise in the individuals with auditory dys-synchrony. Psychophysical studies have demonstrated poor temporal and spectral processing in participants with auditory dys-synchrony and they attributed this as the reason for poor speech perception (Rance et al., 2004; Starr et al., 2003; Zeng et al., 1999).

Vinay and Moore (2007) reported that their participants with auditory dys-synchrony had poor frequency resolution when compared to individuals with normal hearing. Kumar and Jayaram (2010) reported that the poor speech perception abilities are predominantly due to temporal processing deficit. They also saw a poor correlation between pure tone thresholds and speech perception abilities and concluded that audibility is not a major factor that causes impaired speech perception in individuals with auditory dys-synchrony.

Most of the studies in the literature aimed at relating the impaired speech perception to the deficits in phase locking, frequency resolution and temporal processing. A few other studies have seen the speech perception scores in the presence of noise. There are only a few studies which compared the ability to combine speech information from different frequency regions in individuals with hearing loss.

The ability to perceive speech on the basis of sparse cues that are separated in frequency could be important for speech understanding in noisy backgrounds. For example, when the signal to noise ratio is very low, a listener may not have access to the entire spectrum of a speech target and good performance may depend upon the ability to integrate speech fragments that are separated in frequency (Assmann & Summerfield, 2004).

Many studies in speech perception have used vocoders to simulate the spectral channels of cochlear implants. Shannon, Zeng, Kamath, Wygonski and Ekelid(1995), developed a noise-band vocoder to simulate CI speech processing for normal hearing listeners. They found that high level of speech recognition was possible with as few as four spectral channels of information. This result was obtained with simple sentence materials and in quiet listening conditions.

Mlot, Buss and Hall (2010) studied the development of the ability to combine speech information from different frequency regions. They also studied bandwidth required to achieve a low criterion level of speech identification for two frequency bands. They found that children required more bandwidth to identify the stimulus but their ability to integrate the information was similar to that of adults.

Grant, Tufts and Greenberg (2007) examined the intelligibility of speech filtered into relatively narrow spectral bands for both normal-hearing listeners and listeners with sensorineural hearing impairment. They found that ability to integrate the information across the bands was reduced in listeners with sensorineural hearing impairment compared to normals.

Hall, Buss and Grose (2008) considered bandwidth of speech centered either on 500 or 2500 Hz. They varied the bandwidth adaptively to determine the criterion speech bandwidth required to get a score of 15 to 25%. Speech recognition was assessed for low and high bands presented alone, and for the bands presented together. The speech material consisted of Bamford-Kowal-Bench sentences. There was no apparent relation between the criterion normalized bandwidths at the two center frequencies. There were relatively large individual differences in the bandwidth necessary for criterion performance in the hearing-impaired listeners, with criterion normalized bandwidth ranging from approximately 0.28 to 1.06 Hz at 500 Hz, and from approximately 0.14 to 0.54 Hz at 2500 Hz. The criterion speech bandwidths obtained for the hearing-impaired listeners were broadly similar to those obtained by the normal hearing listeners. They found that listeners with mild-moderate sensorineural hearing loss do not have an essential deficit in the ability to integrate across-frequency speech information as their results were comparable with that of individuals with normal hearing.

As is evident from the literature,there are only a few studies (Grant et al., 2007; Hall et al., 2008; Mlot et al., 2010) which have examined the ability to spectrally integrate information across frequencies. These studies have considered only individuals with cochlear hearing loss, and not individuals with auditory dys-synchrony. It is evident from the literature that individuals with auditory dys-synchrony also have poor speech perception abilities (Zeng et al., 1999; Rance et al., 2004) and also difficulty hearing in noise. So it is all the more important to study how the hearing impaired population combine the different spectral information to understand speech, even in noise.

Most of the studies (Hall et al., 2008; Mlot et al., 2010) have used sentences as stimuli which is more redundant. It would be better to use words which are less redundant in speech perception studies. The present study has used filtered words which makes it more difficult to get the redundant information. It is also seen that there is variability among the results of these studies. Hall et al. (2008) said that individuals with cochlear hearing loss has no difficulty in integrating information across frequencies whereas Grant et al. (2007) found that individuals with cochlear hearing loss has difficulty in integrating information across frequencies. Thus there is a need to study spectral integration abilities in individuals with cochlear hearing loss and also in individuals with auditory dys -synchrony.

The present study aimed at finding a criterion speech bandwidth which is necessary to get a minimum (15 to 25%) speech identification score separately for two center frequencies (500 Hz & 2500 Hz) in individuals with normal hearing, cochlear hearing loss and auditory dys-synchrony. The study further aimed at investigating the spectral integration abilities(improvement in speech identification ability that resulted when both bands were presented simultaneously) in all the three groups. Finally, the study intended to investigate if any relation exists between the spectral integration abilities and the speech identification scores in quiet without any modification to the speech stimulus.

Method

Participants

To achieve the goal, three groups of individuals were considered in the present study. The first group being the control group and the next two groups being the clinical groups. The control group consisted of individuals having normal hearing. Individuals having cochlear pathology formed the first clinical group and second clinical group was formed by individuals having auditory dys-synchrony.

Control Group-Individuals with normal hearing sensitivity: This group consisted of 29 individuals with normal hearing sensitivity in the age range of 18 to 50 yrs with a mean age of 28.12 yrs, matched for age with the participants in the clinical group.

All the participants in the control group had normal hearing sensitivity (pure tone thresholds within 15 dB HL in octave frequencies between 250 Hz to 8000 Hz) in both the ears. Participants had greater than 90% speech identification scores in quiet and more than 60% speech identification scores at 0dB SNR.

Immittance evaluation showed type 'A' tympanogram with the presence of acoustic reflexes. None of them had any history of otological symptoms (ear discharge, ear pain, giddiness, or ototoxicity). They did not have any past or present history of neurological dysfunction that was relevant to the present study. All participants were fluent Kannada speakers and did not have any speech or language problems.

Clinical group 1-Individuals with cochlear hearing loss: Consisted of 12 participants in the age range of 18 to 50 years with a mean age of 30.3 years. The participants had acquired mild or moderate sensory hearing impairment as determined by air and bone conduction pure tone audiometry. The pattern of hearing loss was either flat across frequencies or gradually sloping (increase in threshold of around 5-12 dB per octave and the difference between the highest and lowest threshold being no more than 35 dB) from 250 Hz to 8000 Hz. All of participants had speech identification scores proportional to their degree of hearing loss indicating that the hearing loss was predominantly due to cochlear pathology.

Immittance evaluation showed type 'A' tympanogram with either presence, elevated or absence of acoustic reflexes. All participants had absent DPOAEs suggestive of outer hair cell dysfunction. Click evoked ABR was present (proportional to their degree of hearing loss) at

80 dBnHL with a repetition rate of 11.1 clicks/second. There was no past or present history of neurological dysfunction that was relevant to the present study. All participants were fluent Kannada speakers and did not have any speech or language problems.

Clinical group 2-Individuals with auditory dys-synchrony: Consisted of 17 participants in the age range of 18 to 50 years with a mean age of 25.95 years. All of them had bilateral acquired auditory dys-synchrony, with hearing loss not exceeding moderate degree (PTA of 41-55 dB HL). Their speech identification scores were either disproportionate to their degree and configuration of hearing loss or very poor speech perception in noise (SPIN) scores at 0 dB SNR.

Only those individuals who had speech identification scores more than 30% in quiet at 40 dB SL were selected for the present study as the present study required them to identify filtered words. All participants had absent auditory brainstem response (beyond that was expected with the degree of pure tone hearing threshold) at 80 dBnHL with a repetition rate of 11.1 clicks/second. All the participants had DPOAEs and/or cochlear microphonics present.

These participants had normal tympanometric findings with absent ipsilateral and contralateral acoustic reflexes. No other neurological abnormality was present, which was ruled out by an experienced neurologist. All participants were fluent Kannada speakers and did not have any speech or language problem.

Equipments

GSI 61, GSI-TS, Capella OAE analyzer, Biologic Navigator Pro were used to obtain the hearing thresholds, to check the middle ear functioning, OHC's functioning, and also to check retrocochlear involvement respectively. All the equipments were calibrated as per the standards specified by the manufacturer.

Test Environment

Recording of OAEs and all other audiological evaluations, including tests administered to collect data were carried out in a sound treated room. The ambient noise of the test rooms were within the permissible limits as recommended by ANSI (S3.1, 1999).

Test Procedure

All the participants underwent puretone audiometry, immittance audiometry, OAEs and ABR testing. All those participants who met the criteria were selected for the study. The experiment was carried out in three phases:*Preparation of the stimulus, obtaining the criterion speech bandwidth (ie., the minimum bandwidth required to get 15 to 25% SIS),* and *determining the spectral integration score.*

Phase 1: Preparation of the stimulus

The speech stimuli used in this study was bisyllabic words developed by Sreela and Devi (2009) in Kannada. This test contains four lists, each list having 25 bisyllabic words which are phonemically balanced. All the 25 words in each list are equally difficult. For the present study, all the four lists were taken.

The words were recorded in an acoustically treated room. The words were spoken in conversational style by a female native speaker of Kannada. A unidirectional microphone kept at a distance of 10 cm from the speaker's mouth was used. The sampling rate of 44100 Hz and the resolution of 32 bits were used to record the speech stimuli. Each word was recorded thrice to select the best out of three.

Speech intelligibility rating: These recorded words were judged by five native speakers of Kannada. Only those words having good intelligibility were selected for the study.

*Filtering of the words:*Each word was filtered using Adobe Audition software (version 3). The slope of the filter was 60dB/ octave. All the words from all the lists were passed through a band pass filter having either 500Hz or 2500Hz center frequency. The first one was having a low frequency center frequency of 500 Hz and the second one was having a high frequency center frequency of 2500Hz respectively. These center frequencies were also used by Hall, Buss and Grose (2008) in their study on spectral integration. They had selected these center frequencies based on the rationale that frequency components in a speech spectrum predominantly lay between 300 to 3000 Hz. Thus, if a center frequency of 500 Hz and 2500 Hz are taken, these would lie at low and high portions of the speech spectrum respectively. This helps in finding the spectral integration across the speech spectrum. Each word list was filtered using two center frequency having different bandwidths.

Bandwidths considered: The number of bandwidths available for the two center frequencies were different. The words were first passed through a band pass filter with a 500 Hz center frequency. The bandwidth of the filter having 500 Hz as the center frequency was varied from 100 Hz till 1000 Hz in 100 Hz steps. Similarly the bandwidth of filter having 2500 Hz center frequency was also varied from 100 Hz till 3000 Hz in 100 Hz steps.

Initially a pilot study was done on 5 native speakers of Kannada. Initially, filtered speech materials having either low center frequency or higher center frequency were presented to the participants, with the minimum bandwidth. Gradually the bandwidth of the filtered speech was increased. The minimum bandwidth at which the individuals obtained 15 to 25% speech identification scores was noted. This is called as criterion speech bandwidth (CSB) as suggested by Hall et al. (2008).

In the pilot study it was seen that filtered words having 500 Hz center frequency with bandwidths of 100 or 200 Hz was not sufficient for individuals with normal hearing to achieve the criterion score of 15 to 25%. Thus, these two bandwidths were not considered for the study. Similarly for the filtered words having 2500 Hz center frequency bandwidths till 1100 Hz was not sufficient for normal hearing individuals to achieve the criterion score of 15 to 25%. Thus, bandwidths till 1100 Hz were not considered in the study. Table 1 shows the details of bandwidths of two different center frequencies considered for the study.

Maximum bandwidth considered for 500 Hz and 2500 Hz center frequency was 1000 Hz and 3000 Hz respectively. This was not increased further because it would lead to overlapping of bandwidths. For example, if 3100 Hz was considered it would contain frequency components between 950 to 4050 Hz and this will overlap with the 500 Hz center frequency having a bandwidth of 1000 Hz (0-1000 Hz).

Table 1: Bandwidths used for the study having two different center frequencies

500 Hz center frequency		2500 Hz center frequency	
Frequency range (Hz)	Bandwidth (Hz)	Frequency range (Hz)	Bandwidth (Hz)
350-650Hz	300Hz	1900-3100	1200
300-700Hz	400Hz	1850-3150	1300
250-750Hz	500Hz	1800-3200	1400
200-800Hz	600Hz	1750-3250	1500
150-850Hz	700Hz	1700-3300	1600
100-900Hz	800Hz	1650-3350	1700
50-950Hz	900Hz	1600-3400	1800
0-1000Hz	1000Hz	1550-3450	1900
-	-	1500-3500	2000
-	-	1450-3550	2100
-	-	1400-3600	2200
-	-	1350-3650	2300
-	-	1300-3700	2400
-	-	1250-3750	2500
-	-	1200-3800	2600
-	-	1150-3850	2700
-	-	1100-3900	2800
-	-	1050-3950	2900
-	-	1000-4000	3000

Phase 2: Obtaining Criterion Speech Bandwidth (CSB)

Criterion Speech Bandwidth was established using two steps. Step one was to obtain initial bandwidth for CSB and the second step to establish the CSB.

Step to obtain Initial Bandwidth for CSB: To obtain the initial level the stimuli were presented through a calibrated 2 channel diagnostic audiometer GSI-61 with TDH 50P earphones. Presentation level was kept at 40 dB SL for all the participants and it was monitored through audiometer. Responses were obtained from the participants by instructing them either to repeat or write the words. Participants were instructed to guess the words if it was not clearly perceived. Only one ear was considered for all the participants to reduce the practice effect. The ear which fulfilled the criteria was selected for testing. If both the ears of a single subject passed the criteria then their right ear was considered for testing. Experimenter didn't give any feedback regarding their responses during the testing.

With the goal of predicting the CSB, filtered words were presented to the participants. At first filtered words having center frequency of 500 Hz were presented. An initially filtered word with largest bandwidth of 1000 Hz was presented for familiarization. Two filtered words were presented at each bandwidth. If the participants failed to identify both the words, then bandwidth was increased by 100 Hz and the next set of filtered words were presented. For example, in the Table 2, after familiarizing the participants by presenting filtered word with largest bandwidth, filtered words having a bandwidth of 300Hz were presented. Since the participant could not identify both the words at this bandwidth, the bandwidth was increased by 100 Hz ie, 400Hz and again two filtered words were presented. When the participants were able to identify both the words at a particular bandwidth, this was considered as initiation bandwidth for CSB. In the Table 2, at the bandwidth of 500 Hz, the

Table 2: Procedure to obtain initiation bandwidth for CSB

Bandwidth (center frequency 500 Hz)	Response(Word identification-two words at each band width)	
	1st word	2nd word
1000 Hz, for familiarization	present	present
300 Hz	absent	absent
400 Hz	absent	absent
500 Hz	present	present

participant correctly identified the filtered words. Thus the initiation bandwidth for CSB is 500Hz.

The same procedure was also followed for the 2500 Hz center frequency to obtain the initiation bandwidth for CSB. This procedure was followed to minimize the presentation of full list to obtain CSB.

Phase 3: Step to obtain CSB

Criterion speech bandwidth was the minimum bandwidth required to get 15 to 25% word identification scores. Thus, in the next step of the study a full list of 25 filtered words were presented to the participants at their initiation bandwidth for CSB's for both the center frequencies to see whether it could give the criterion score of 15 to 25%. Each correct word was given a score of 4%, thus 25 words in a list makes a total of 100%. Hall et al. (2008) also considered criterion score of 15 to 25%. In case they failed to obtain 15 to 25% score at their initiation bandwidth for CSB then the bandwidth was increased at the order of 100 Hz and again a full list of 25 filtered words was presented. Bandwidths were increased till the criterion score was achieved. The bandwidth at which the score of 15 to 25% was obtained was considered as the CSB.

The relatively low criterion of 15 to 25 % was considered to ensure that performance is below 100% when both the bands are presented together.

Determining the Spectral integration: Two CSBs were obtained for all the participants. The first one for the low center frequency of 500 Hz and the second one for the higher center frequency of 2500 Hz. To determine the spectral integration abilities, words having both the CSB's were presented to the participants. A full list of 25 words was used and the word identification scores were calculated.

Results

For each participant, CSB's for 500 Hz center frequency and 2500 Hz center frequency was noted. These CSB's were then divided by their respective center frequencies to obtain normalized CSB. Speech identification scores were obtained by presenting words having both the CSBs. These values were taken for comparison across groups.

Within group statistical analyses were done for comparing the parameters within the same group. Paired t-test was carried out to determine whether a significant difference existed between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at low and high center frequencies, within each group. A Pearson correlation was done to see the correlation between the spectral integration scores and the speech identification scores obtained in quiet without any modification to the speech stimulus.

Between group statistical analyses were done to compare parameters across the groups. Independent t- test was carried out to see the group differences for normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500 Hz, between individuals with normal hearing and individuals with cochlear hearing loss. One way ANOVA was not done for group comparison of normalized bandwidth at 500 Hz because individuals with auditory dys-synchrony could not get minimum speech identification scores even at the maximum bandwidth at the 500 Hz center frequency used in the study. One way ANOVA was done to see whether a significant difference existed between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 2500 Hz, across the three groups .One way ANOVA was also done to see whether a significant difference existed between spectral integration scores obtained across the groups. Duncan's post hoc analysis was done to see the pairwise differences when the ANOVA results were significant.

Individuals with Normal Hearing

The mean and the standard deviation for the normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500 Hz, 2500 Hz center frequency and for the speech integration scores were calculated for all the 29 individuals with normal hearing sensitivity. Details are given in Table 3.

From the table it can be seen that the normalized minimum bandwidth achieving minimum speech identification scores at 500 Hz center frequency was greater than the bandwidth required at 2500 Hz center frequency. Paired t- test was carried out to determine whether a significant difference existed between normalized bandwidth at these two center frequencies. Results showed that there was a significant difference [t= (3.73), 28 $p<0.001$] between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at low and high center frequencies.

Table 3: Mean, Standard Deviation (SD), minimum and maximum values for the normalized criterion speech bandwidths at two different center frequencies and also speech integration scores obtained in individuals with normal hearing sensitivity

	Mean (kHz)	SD	Min	Max
Normalized CSB at 500 Hz	0.76 (N=29)	0.2	0.68	0.83
Normalized CSB at 2500 Hz	0.61 (N=29)	0.08	0.58	0.64
Spectral integration	90.34% (N=29)	3.30	89.09%	91.50%

Correlation between the spectral integration scores and speech identification scores obtained in quiet without any modification to the speech stimulus was not done in this group, since all the participants in the group got 100% speech identification scores in quiet without any modification to the speech stimulus.

Individuals with Cochlear Hearing Loss

The mean and the standard deviation for the normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500 Hz, 2500 Hz center frequency and for the speech integration scores were calculated for all the 12 individuals with cochlear hearing loss. Details are given in Table 4.

Table 4 shows that normalized bandwidth required for 500 Hz is more than that required for 2500 Hz center frequency. Paired t- test was carried out to determine whether a significant difference existed between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at low and high center frequencies. Results showed that there was no significant difference (t= (1.96), 11 p>0.05) between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at low and high center frequencies.

Pearson correlation was done to see the relationship between the spectral integration scores and the speech identification scores obtained in quiet without any modification to the speech stimulus in individuals with cochlear hearing loss. Results of the correlational analysis showed that there was no significant correlation between the spectral integration scores and the speech identification scores obtained in quiet without any modification to the speech stimulus in individuals with cochlear hearing loss (r=0.35, p> 0.05).

Individuals with Auditory Dys-synchrony

Mean for the normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500 Hz was not calculated as none of the individuals with auditory dys-synchrony could get a minimum speech identification score of 15 to 25%, even at the maximum bandwidth of 500Hz center frequency used in the study.

Table 4: Mean, Standard Deviation (SD), minimum and maximum values for the normalized criterion speech bandwidths at two different center frequencies and for the speech integration scores obtained in individuals with cochlear hearing loss

	Mean (kHz)	SD	Min	Max
Normalized CSB at 500 Hz	0.92 (N=12)	0.29	0.73	1.1
Normalized CSB at 2500 Hz	0.74 (N=12)	0.17	0.63	0.84
Spectral integration scores	92.3% (N=12)	3.17	90.31%	94.84%

Table 5: Mean, Standard Deviation (SD), minimum and maximum values for the normalized criterion speech bandwidths at 2500 Hz center frequency and also speech integration scores obtained in individuals with auditory dys-synchrony

	Mean (kHz)	SD	Min	Max
Normalized CSB at 2500 Hz	1.08 (N=17)	0.15	1	1.16
Spectral integration scores	30.11% (N=17)	13.71	23.06%	37.17%

The mean and the standard deviation for the normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) at 2500 Hz and for the speech integration scores were calculated for all the 17 individuals with auditory dys-synchrony. Spectral integration scores were calculated by presenting filtered words having both CSBs (500 Hz and 2500 Hz center frequency). As none of the individuals with auditory dys-synchrony could get a minimum speech identification scores even at the maximum bandwidth of 500 Hz center frequency, for calculating spectral integration scores maximum bandwidth at 500 Hz center frequency was presented along with the CSB obtained at 2500 Hz center frequency. Details are given in Table 5.

There were 2 different patterns of integration seen in these individuals. This included negative spectral integration and poor spectral integration (marginal im-

Table 6: Scores/ Criterion scores obtained at 500 Hz and 2500 Hz center frequencies and also the spectral integration scores in individuals with auditory dys-synchrony

Participants	500 Hz center frequency with max bandwidth of 1000 Hz	Crit. score at 2500 Hz	Spectral integrat-ion score	SI scores in quiet with-out stimulus modification
1	4%	24%	12%	32%
2	8%	24%	16%	32%
3	0%	20%	16%	36%
4	0%	24%	4%	36%
5	0%	20%	16%	36%
6	8%	20%	28%	32%
7	8%	24%	40%	88%
8	0%	20%	32%	88%
9	8%	24%	40%	60%
10	8%	20%	32%	40%
11	4%	24%	52%	88%
12	0%	20%	52%	80%
13	0%	24%	40%	92%
14	0%	20%	24%	68%
15	12%	16%	36%	100%
16	0%	20%	36%	76%
17	0%	20%	32%	76%

provement). Negative spectral integration means when both the low and high center frequency bands were presented together to the participants instead of getting a better integrated score by combining the information in both the bands, these individuals got a poorer score than the criterion score obtained at 2500 Hz center frequency CSB. Out of the 17 individuals with auditory dys-synchrony 5 had negative spectral integration. The remaining 12 individuals had less advantage of spectral integration (marginal improvement). The details are given in Table 6.

From the Table 6, it is evident that none of the individuals could achieve a criterion score of 15 to 25% at the 500 Hz center frequency. All of them achieved a criterion score at 2500 Hz center frequency. When the information in both the bands was presented together first 5 participants got poorer scores, even poorer than their criterion scores obtained at 2500 Hz center frequency indicating a negative spectral integration. All these five participants had poor speech identification scores in quiet without any modification made in the speech stimulus with their scores ranging from 32% to 36%.

The remaining 12 participants with auditory dys-synchrony got better spectral integration values when compared to the first five participants with the scores ranging from 28% to 52%. Among the 12 participants, 10 had speech identification scores in quiet of 60% or above. Only the participants 6 and 10 had speech identification scores less than 50% in this group.

Pearson correlation was done to see the relationship between the spectral integration scores and the speech identification scores in quiet without any modification to the speech stimulus. The results showed that there was a significant positive correlation between the spectral integration scores and the speech identification scores obtained in quiet without any modification to the speech stimulus ($r=0.641$, $p<0.01$). This means that, better the Speech identification scores in quiet without any modification made in the speech stimulus, better the spectral integration scores and vice-versa.

Across Group Comparisons

Mean, Standard Deviation of normalized criterion speech bandwidths at 500 Hz and 2500 Hz center frequencies were compared across the groups. The results are given in the Figure 1.

None of the individuals with auditory dys-synchrony could achieve a criterion score even at maximum bandwidth at 500 Hz center frequency. It is seen that individuals with normal hearing obtained the criterion scores with least CSBs at both 500 Hz and 2500 Hz center frequencies followed by individuals with cochlear hearing loss and then the individuals with auditory dys-synchrony (CSB at 2500 Hz). The variability was relatively great among the individuals with cochlear hearing loss for the CSBs at both 500 Hz and 2500 Hz center frequencies. The mean and standard deviation for spectral integration scores were also compared across the groups. The details are given in Table 7.

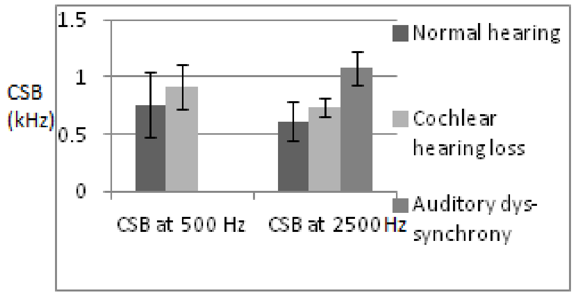


Figure 1: Mean, Standard Deviation (SD) at two different center frequencies obtained across all the three groups.

Table 7:

When the spectral integration scores were compared across the groups it was seen that both normal hearing individuals and individuals with cochlear hearing loss performed almost equally. Individuals with auditory dys-synchrony had very less spectral integration scores compared to the other two groups and also the variability was more in this group which is evident from the larger standard deviation value.

Normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500Hz was compared across two groups (between individuals with normal hearing and those with cochlear hearing loss) since the individuals with auditory dys-synchrony could not achieve the criterion score even at the maximum band width of 500 Hz center frequency. Hence, at 500 Hz bandwidth an independent t-test was used to compare the normalized minimum bandwidth required to achieve minimum speech identification scores (normalized CSB) across individuals with normal hearing and those with cochlear hearing loss. Results showed that there was no significant difference in normalized minimum bandwidth required to achieve minimum speech identification scores at 500Hz center frequency between the two groups ($t=2$, $p>0.05$).

One way ANOVA was done to see whether a significant difference existed between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 2500 Hz, across the three groups. Results showed that there was a significant difference across groups [$F(2,55)=77.4$ $p<0.001$]. Duncans post hoc analysis was done to see if all the three groups differed significantly from each other for the normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 2500 Hz. It was found that all the three groups differed significantly from each other ($p<0.05$).

One way ANOVA was done to see whether a significant difference is present across the groups for the spectral integration scores. It was found that there was a significant difference [$F(2,55)=356.86$, $p<0.001$] across the groups. Duncans post hoc analysis was done to see if all the three groups differed significantly from each other for spectral integration scores. It was found that individuals with auditory dys-synchrony were significantly different from the other two groups in terms of spectral integration scores ($p<0.05$).

Table 7: Mean, Standard Deviation (SD) for the speech integration scores obtained across all the three groups.

Groups	Mean	SD
NH	90.34%	3.30
CH	92.3%	3.17
AD	30.11%	13.71

Discussion

Results showed that, for individuals with normal hearing the normalized CSB at 500Hz center frequency ranged from 0.68 to 0.83 and for individuals with cochlear hearing loss it ranged from 0.73 to 1.1. This is larger in comparison to the previous studies. Hall, Buss and Grose (2008) in their study said that for individuals with normal hearing the criterion normalized bandwidth at 500 Hz center frequency ranged from 0.27 to 0.57 and for individuals with cochlear hearing loss it ranged from 0.28 to 1.06. The difference in the present study from the study by Hall et al. (2008) might be due to the type of stimuli used. They used filtered sentences whereas, in the present study filtered words was used as stimuli and this is probably because sentences are more redundant than words.

There was no significant difference in normalized bandwidth required to achieve minimum speech identification scores at 500 Hz center frequency between individuals with normal hearing and individuals with cochlear hearing loss. However there was more variability in individuals with cochlear hearing loss. Similar results were also discussed by Hall et al. (2008). This can be explained with the degree and pattern of hearing loss considered in the present study. The current study has taken only individuals with flat or gradually sloping hearing loss of mild-moderate degree. Glasberg and Moore (1989) said that individuals with cochlear hearing loss of only more than moderate degree have major problems with frequency resolution. Thus, most of the participants in cochlear hearing loss group would not have had a problem with their frequency resolution and temporal coding that much which could bring a significant difference between individuals with normal hearing and individuals with cochlear hearing loss.

Individuals with auditory dys-synchrony could not achieve the criterion score even at the maximum bandwidth at 500 Hz center frequency. This can be attributed to their poor frequency resolution at low frequencies (Zeng & Liu, 2006) due to which their speech perception was severely affected at low center frequency.

Results showed that, for individuals with normal hearing the CSB at 2500 Hz ranged from 0.58 to 0.64 and for individuals with cochlear hearing loss it ranged from 0.63 to 0.84 and in individuals with auditory dys-synchrony it ranged from 1 to 1.16. The bandwidth required by individuals with normal hearing and also individuals with cochlear hearing loss at both the center frequencies in the current study were larger in comparison to the previous studies. Hall et al. (2008), in their study said that for individuals with normal hearing the criterion normalized bandwidth at 2500 Hz center frequency ranged from 0.22 to 0.48 and for individuals with cochlear hearing loss it ranged from 0.14 to 0.54. Mlot, Buss and Hall (2010) have also reported similar results as that of Hall et al. (2008). The difference in the present study from the previous studies can be again explained by the type of stimuli used.

All the three groups differed significantly from each other for the normalized bandwidth required to achieve minimum speech identification scores having 2500 Hz center frequency. This can be explained with the explanation given by Lorenzi, Gilbert, Cam, Gamier and Moore (2006) who reported that individuals with cochlear hearing loss has difficulty using the fine structure cues which are of high frequency information. So speech processing varies based on the frequency resolution at a particular frequency and also it varies across listeners. Thus in the present study, individuals with cochlear hearing loss would have had poorer frequency resolution at high center frequency due to which they required wider CSB than that of normal hearing individuals. Whereas, individuals with auditory dys-synchrony required the widest band width among the three groups to achieve minimum speech identification scores (normalized CSB) at 2500 Hz center frequency. Though temporal processing is majorly affected in these individuals they also have spectral processing difficulties (Zeng , Oba, Garde, Sininger & Starr, 2001; Rance, McKay & Grayden, 2004; Starr et al., 2003). Vinay and Moore (2007) reported poor ability in individuals with auditory dys-synchrony to detect tones in presence of noise and they also attributed this to the poor phase locking in these individuals. Therefore all these reasons would have contributed for poorer performance in this group.

Results also showed a significant difference between normalized bandwidth required to achieve minimum speech identification scores at 500 Hz and 2500 Hz center frequencies in individuals with normal hearing. This finding is in accordance with the study done by Mlot, Buss and Hall (2010) where they found that normalized CSB was significantly smaller for the band centered on 2500 Hz than that for the band centered on 500 Hz. This result can be explained with frequency band importance. The greater importance of the higher frequency band may explain the fact that it carries more information essential for determining consonant place, which is more essential in enabling the listener to discriminate among words (Kasturi, Loizou, Dorman & Spahr, 2002).

In individuals with cochlear hearing loss there was no significant difference between normalized bandwidth required to achieve minimum speech identification scores at 500 Hz and 2500 Hz center frequencies. This result is in contrary to the results discussed by Hall et al. (2008). He found that even individuals with cochlear hearing loss require lesser CSB at 2500 Hz center frequency compared to 500 Hz center frequency. In their study they had given a high frequency boost to the high frequency band to ensure the constant audibility and also to reduce the effect of upward spread of masking among hearing impaired listeners, which was not done in the present study. Another reason might be the type of the stimuli used in both the studies. Hall et al. (2008), had used filtered sentences where as the present study used filtered words as stimuli.

Comparison between normalized bandwidth required to achieve minimum speech identification scores (normalized CSB) at 500 Hz and 2500 Hz center frequencies was not made in the group with auditory dys-synchrony as none of them could achieve the criterion score even at the maximum bandwidth at 500 Hz center frequency.

Results showed that individuals with normal hearing and individuals with cochlear hearing loss performed similar in spectral integration scores (with both groups having more than 90% scores when both bands were presented together). Similar findings have been reported in individuals with normal hearing and individuals with cochlear hearing loss by Hall et al. (2008). However the amount to which integration of the information occurred was different in the present study in individuals with normal hearing and individuals with cochlear hearing loss. Hall et al. (2008) in their study found that when the individual band which gives a criterion score of 15-25% were presented together spectral integration scores were better than 70%. Resuts of Mlot et al. (2010) also closely agrees with that of Hall et al. (2008) finding. In the present study when the low and the high frequency bands were presented together both the individuals with normal hearing and individuals with cochlear hearing loss got spectral integration scores of more than 90%. In their studies they obtained criterion score of 15-25% at smaller CSBs than that of the present study. In the study by Hall et al. (2008) the CSBs for low and high center frequencies were 0.41 and 0.35 respectively for normal hearing adults. On contrary in the present study CSBs for low and high center frequencies were 0.76 and 0.61 respectively for normal hearing adults. Thus when a two large spectral bands are presented together integration occurs across many frequencies than when smaller bands are presented together. This might have resulted in better integration scores of more than 90%.

Individuals with auditory dys-synchrony had significantly poor spectral integration scores than individuals with normal hearing and individuals with cochlear hearing loss.This can be explained based on the degree of dys-synchrony in these individuals. It is evident from the literature that these individuals have poor phase locking abilities which results in poor pitch processing mainly at low frequencies (Zeng & Liu, 2006). Thus those five individuals who had negative spectral integration would have had very poor pitch processing at low frequencies to the extent that it even interrupted their processing of high frequency information when both the CSBs were presented together. In other words they have failed to utilize the information at and around 500 Hz center frequency, rather the energy of this level would have caused upward spread of masking leading to the masking of high frequency signal which resulted in reduced performance. Individuals with auditory dys-synchrony also shows excessive masking effect (Zeng, Kong, Michalewski & Starr, 2005) which would further enhance the upward spread of masking and this would have resulted in poorer spectral integration scores, even poorer than their criterion scores obtained at 2500 Hz center frequency when the information in both the bands was presented together. This can be further supported by the fact that all the five participants had poor speech identification scores (32% - 36%) in quiet without any modification made in the speech stimulus.

Remaining 12 participants had poor spectral integration. Both the individuals with normal hearing and individuals with cochlear hearing loss, the spectral integration scores were greater than 90%, where as in individuals with auditory dys-synchrony the spectral integration scores ranged from 28%-52%.

The reason for poor performance compared to other two groups can be again explained using the poor pitch processing in individuals with auditory dys-synchrony. Reduced pitch processing in individuals with auditory dys-synchrony limits them from combining the information across the frequency bands effectively as in case of individuals with normal hearing and also of cochlear hearing loss. However these 12 individuals got better spectral integration scores compared to the other 5 individuals with auditory dys-synchrony. This might be because the degree of dys-synchrony was milder in this group. This is supported by the fact that 10 individuals among the 12 individuals with auditory dys-synchrony had their speech identification scores greater than 60% in quiet without any modification made in the speech stimulus, which suggests lesser degree of dys-synchrony. Results showed no correlation between the spectral integration scores and speech identification scores obtained in quiet without any modification made in the speech stimulus in individuals with cochlear hearing loss.

A positive significant correlation between the spectral integration scores and speech identification scores obtained in quiet without any modification made in the speech stimulus was seen in individuals with auditory dys-synchrony. In individuals with auditory dys-synchrony, only those individuals who had good speech identification had better spectral integration scores. This can be explained based on the frequency resolution at low frequencies. Those individuals who had better frequency resolution could obtain better speech identification scores in quiet which in turn resulted in improved ability to combine information across frequency bands.

Conclusion

These findings of the study are helpful while selecting hearing aid features for these individuals. Most of the individuals with moderate sensorineural hearing loss of flat or slightly sloping pattern will benefit from multi channel hearing aids as they have very good ability to combine the information across the frequencies. In individuals with auditory dys-synchrony it is better to select a hearing aid with lesser number of channels as they already have very poor abilities to combine information across the frequencies. It is also best to give them a hearing aid with best noise reduction strategies which will help to remove noise which are mainly of low frequencies. Even while prescribing them channel specific gain it is wise to give lesser gain at low frequencies to reduce the upward spread of masking, which can cause deleterious effect, as seen in the present study.

This study can be used as a tool to study the spectral integration abilities in different clinical groups. This can be used as a tool to assess the speech perception abilities in difficult listening situations as we are using filtered words. This study can also be used to differentiate between individuals with cochlear hearing loss and those with auditory dys-synchrony.This can be used to explain physiological basis for the speech perception abilities of different clinical groups to some extent. Further studies on CSBs required for speech perception may assist us in selection of hearing aids by helping us decide about the optimum number of channels required for each individual.

References

American National Standards Institute. (1999). *Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms, ANSI S3.1-1999*, New York: American National Standards Institute.

Assmann, P. F. & Summerfield, A. Q. (2004). The perception of speech under adverse conditions.

In: S. Greenberg, W.A. Ainsworth, A.N. Popper and R.R. Fay (Eds.), *Speech Processing in the Auditory System, 14* (pp. 231-308). New York City: Springer Handbook of Auditory Research.

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in participants with unilateral and bilateral cochlear impairments. *Journal of the Acoustical Society of America, 79*, 1020-1033.

Glasberg, B. R., & Moore, B. C. J. (1989). Difference limens for phase in normal and hearing-impaired participants.*Journal of the Acoustical Society of America, 86,* 1351-1365.

Grant, K. W., Tufts, J. B., & Greenberg, S. (2007). Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impaired individuals. *Journal of the Acoustical Society of America, 121,* 1164-1176.

Hall, J. W., Buss, E., & Grose, J. H. (2008). Spectral integration of speech bands in normal hearing and hearing impaired listeners. *Journal of the Acoustical Society of America, 124,* 1105-1115.

Hamernik, R. P., & Davis, R. I. (1988). *Noise and Hearing Impairment.* In B.S. Levy., & D.H. Wegman (Ed.), *Occupational Health, (*247-261). Boston, Little, Brown and Co.

Kasturi, K., Loizou, P. C., Dorman, M., & Spahr, T. (2002). The intelligibility of speech with 'holes' in the spectrum. *Journal of the Acoustical Society of America,112,* 1102-1111.

Kumar, U. A., & Jayaram, M. (2010). Speech perception in individuals with auditory dys-synchrony. *The Journal of Laryngology & Otology*, *125,* 236-245.

Lorenzi, C., Gilbert, G., Cam, H., Gamier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy Sciences, 103,* 18866-18869.

Mlot, S., Buss, E., & Hall, J. W. (2010). Spectral Integration and Bandwidth Effects on Speech Recognition in School Aged Children and Adults. *Ear and hearing, 31,* 56-62.

Plomp, R. (1994). Noise, amplification, and compression: Considerations of three main issues in hearing aid design. *Ear and Hearing, 15,* 2-12.

Rance, G. (2005). Auditory dys-synchrony/dys-synchrony and it's perceptual consequences. *Trends in Amplification, 9,* 1- 43.

Rance, G., Cone-Wesson, B., Wunderlich, J., & Dowell, R. (2002). Speech perception and cortical event related potentials in children with auditory dys-synchrony. *Ear and Hearing, 23,* 239-253.

Rance, G., McKay, C., & Grayden, D. (2004). Perceptual characterization of children with auditory dys-synchrony. *Ear and Hearing, 25,* 34-46.

Rapin, I., & Gravel, J. (2003). "Auditory dys-synchrony": physiologic and pathologic evidence calls for more diagnostic specificity. *International Jouranal of Pediatric Otorhinolaryngology 67,* 707-728.

Shannon, R.V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995) Speech recognition with primarily temporal cues. *Science, 4*, 270-303.

Starr, A., Michalewski, H. J., Zeng, F. G., Brooks, S. F., Linthicum, F., Kim, C. S., et al. (2003). Pathology and physiology of auditory dys-synchrony with a novel mutation in the MPZ gene**.** *Brain*, *126,* 1604-1619.

Starr, A., Picton, T. W., Sininger, Y., Hood, L., & Berlin, C. I. (1996). Auditory dys-synchrony. *Brain, 119,* 741-753.

Sreela, P. K., & Devi, N. (2009). *Effect of reverberation on speech identification using hearing aids*. Unpublished master's thesis based article, University of Mysore, Mysore, India.

Thornton, A. R., & Abbas, P. J. (1980). Low-frequency hearing loss: perception of filtered speech, psychophysical tuning curves, and masking. *Journal of the Acoustical Society of America, 67,* 638-643.

Tyler, R. S. (1986). Frequency resolution in hearing-impaired listeners. In: Moore, B.C. J. (Ed.), *Frequency Selectivity in Hearing*. London, Academic Press.

Vinay, & Moore, B. C. J. (2007). Ten (HL)-test results and psychophysical tuning curves for participants with auditory dys-synchrony.*International Journal of Audiology, 46,* 39-46.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology, 93,* 3050-3063.

Zeng, F. G., & Liu, S. (2006). Speech perception in auditory dys-synchrony participants. *Journal of Speech & Hearing Research, 42*(2), 367-380.

Zeng, F. G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (2001). Psychoacoustics and speech perception in auditory dys-synchrony. In: Y. Sininger, & A. Starr (Eds.), *Auditory dys-synchrony: A new perspective on hearing disorder,* (141-164): Singular publishing group, Canada.

Zeng, F. G., Oba, S., Sininger, Y. S., & Starr, A. (1999). Temporal and speech processing deficits in auditory dys-synchrony. *Neuroreport*, *10,* 3429-3435.

Perception of Spectrally Enhanced Speech through Companding in Individuals with Auditory Dys-synchrony

[1]Shachi Vasishta,[2]Animesh Barman

Abstract

The study was taken up with the aim to know whether spectrally enhanced speech through compounding could improve speech perception of individuals with Auditory dys-synchrony in quiet and at different signal to noise ratios. To achieve the objective, 15 ears of fifteen individuals with auditory dys-synchrony and 15 ears of fifteen normal hearing individuals (mean age 27 years) were taken. Speech identification abilities of these individuals were assessed in quiet and at three different SNRs (+15 dB, +10 dB, +0 dB) for VCV syllables with and without companding. The minimum SNR at which these individuals correctly identified 50% of the words in a sentence was also assessed using Quick SIN sentences test in Kannada with and without companding. Results showed that as the SNR increased, the number of syllables correctly identified increased for VCV syllables in both with and without companding conditions , in both the groups. Normal hearing individuals showed significantly better performance for companded stimuli compared to non companded VCV syllables at + 0 dB SNR. Individuals with auditory dys-synchrony demonstrated this trend in quiet. Also, they required significantly lower SNR (better) for companded sentences than non companded sentences to correctly identify 50% words in a sentence. For companded sentences, SNR required to correctly identify 50% of the words in a sentence was lower than that of the non companded sentences, in both the groups. However, the normal hearing individuals performed well at a significantly lower (better) SNR than the individuals with auditory dys-synchrony for both companded and non companded test materials.

Keywords: *Companding, Auditory Dys-synchrony, Spectral enhancement, speech perception*

Introduction

Auditory dys-synchrony (AD) is a distinct hearing disorder characterized by auditory nerve dysfunction in the presence of normal outer hair cell activity (Starr, Picton, Sininger, Hood & Berlin, 1996). Desynchronized discharges at the level of 8th nerve and brainstem have been proposed as one of the underlying pathophysiologic mechanisms (Zeng, Oba, Garde, Sininger & Starr 1999; Kraus et al., 2000; Kumar & Jayaram, 2005). Psychophysical studies indicated that the consequences of disrupted auditory nerve activity are reflected as a significant impairment in temporal processing and difficulty in understanding speech that is disproptionate to the degree of hearing loss measured by pure tone thresholds (Siniger, & Oba, 2001; Zeng et al., 1999).

Difficulty in understanding speech, particularly in noise, is found to be a consistent problem reported by individuals with AD. Studies have investigated speech perception in noise in individuals with AD and illustrate that the noise has more detrimental effect on speech perception than that observed for listeners with normal hearing and those with cochlear hearing loss (Rance et al., 2007; Zeng & Liu, 2006).

The psychoacoustical studies have demonstrated impaired frequency discrimination for low frequency sounds (500 Hz) and with near normal values at 4000 Hz (Zeng et al. 1999; Rance, Mckay & Grayden 2004). Consistent with these findings, it has been observed that individuals with AD show good identification for phonemes that lie in the high frequency range than those phonemes that lie in the low frequency range (Rance & Barker, 2008; Narne & Vanaja, 2008).

Individuals with AD show severely affected temporal processing abilities that seem to be the basis of their poor speech perception (Zeng et al., 1999). Psychophysical measures showed that the disrupted neural activity significantly impairs timing related perception, such as pitch discrimination at low frequencies, temporal integration, gap detection, temporal modulation detection, forward and backward masking, binaural beats, signal detection in noise, and sound localization using interaural time differences (Zeng, Kong, Michalewski & Starr, 2005). Listeners with AD typically required silent periods of 20 ms or more to detect a gap compared to less than 5 ms in normal listeners (Rance et al, 2008). Also, speech signals are often degraded by noise, in real life situations. While normal-hearing listeners are capable of extracting the critical information from noisy speech, this ability is affected in individuals with AD.

Management of AD continues to be difficult and challenging. Persons with cochlear hearing loss derive significant benefit from hearing aids which employ nonlinear compression circuits. All these hearing aids assume abnormal functioning of outer hair cells (Berlin, Hood, Hurley & Wen 1996). Hence, these aids are of not much use for individuals with auditory dys-synchrony who

[1]Email: shachi.vasishtha@gmail.com,
[2]Reader in Audiology, Email: nishiprerna@yahoo.com

have normal outer hair cell functioning as supported by Rance, Cone-Wesson, Wunderlich and Dowell, (2002). Several other management strategies which can be used with individuals with auditory dys-synchrony include FM systems, cochlear implants, perceptual training, speech reading and cued speech (Kraus, 2001). Many studies have emphasized cochlear implants as the treatment of choice for AN (Rance & Barker, 2008). However, the invasive nature of cochlear implants and their doubtful efficacy points to a need for research on alternative strategies to improve speech intelligibility. This is particularly true for individuals with mild AD (Zeng & Liu, 2006).

Turicchia and Sarpeshkar (2005) have proposed a novel spectral enhancement scheme, companding, which combines two-tone suppression and dynamic gain control to increase the spectral contrast. Studies have shown that companding is also present along the auditory pathway. Both cochlea and the cochlear nucleus perform logarithmic compression on the input signals, while the brain performs exponential expansion (Zeng & Shannon, 1999). Considering this hypothesis, a signal processing strategy has been developed where certain signal will be compressed and certain frequencies will be enhanced which is termed as companding. Implementing the companding strategy, Bhattacharya and Zeng (2007) showed significant improvement in both phoneme and sentence perception in noise, in the cochlear implants users. However, the usefulness of this option has not been investigated in individuals with AD.

It is thus essential to study if listeners with AD can benefit from spectral enhancement of speech through companding. A systematic comparison with normal hearing listeners at various SNRs will be more appropriate to determine if the effect is level dependent or not. Keeping all this in mind the present study aimed to a) know whether companding of speech stimuli helps to improve speech intelligibility in individuals with normal hearing and those with AD, b) identify benefit of spectrally enhanced speech through companding at different SNRs in both the groups and c) determine how individuals with normal differ from those with AD in their performance for companded stimuli at quiet and at different SNRs.

Method

Participants

The participants were divided into clinical group consisting of those with AD and control group consisting of individuals with normal hearing.

Clinical group: consisted of 15 ears from 10 participants, in the age range of 15 to 42 years (mean age 27 years), fulfilling the criteria of AD in both ears. The demographic data and audiological test findings of 10 participants are given in Table 1. All the participants in the clinical group had degree of hearing loss ranging from mild to moderately severe sensorineural hearing loss. They had acquired, post lingual hearing loss and had disproportionately poor speech identification scores in relation to their pure tone threshold or poor SIS at 0 dB SNR. Further, they had absent auditory brainstem responses beyond that can be expected from thei degree of hearing loss, but had present otoacoustic emissions indicating normal OHC function. They had type 'A' tympanogram with no ipsi or contralateral reflexes.

Table 1: The Audiological test findings in individuals with AD

Participants	Age/ Gender	Ear	Severity of Hearing loss	SIS in quiet	SIS at 0 dB SNR	Pure tone configuration
AD1	42/ F	Right	Mild	34%	0%	Reverse slope
		Left	Mild	32%	0%	Reverse slope
AD2	32/M	Right	Moderately severe	88%	0%	Reverse slope
		Left	Moderately Severe	60%	0%	Reverse slope
AD3	25/F	Right	Minimal	92%	24%	Flat
		Left	Minimal	86%	28%	Flat
AD4	15/F	Right	Mild	84%	36%	Sloping
		Left	Mild	88%	44%	Flat
AD5	19/F	Right	Mild	88%	32%	Reverse slope
		Left	Mild	88%	28%	Flat
AD6	29/M	Right	Mild	76%	25%	Trough shape
		Left	Mild	76%	25%	Trough shape
AD7	26/M	Right	Mild	40%	0%	Reverse slope
		Left	Minimal	36%	0%	Reverse slope
AD8	27/ F	Right	Moderate	68%	0%	Reverse slope
		Left	Mild	92%	0%	Reverse slope
AD9	38/M	Right	Mild	88%	32%	Flat
		Left	Mild	100%	36%	Flat
AD10	19/F	Right	Moderately Severe	40%	0%	Flat
		Left	Moderately Severe	44%	0%	Flat

No other neurological symptoms were reported by these participants. Peripheral neuropathy or space-occupying lesion was ruled out by a neurologist. Any other otological disorder including middle ear infections were ruled out by an Otologist. All participants were fluent in Kannada and had no speech and language problems.

Test Environment

All the tests were carried out in a sound treated room. Noise levels in the test room were within permissible limits as per ANSI S3.1-1991.

Test equipment

A calibrated double channel diagnostic audiometer GSI- 61 with TDH- 50 P headphones were used for pure tone and speech audiometry. A calibrated immittance meter GSI tympstar was used to confirm the normal middle ear function through tympanometry and acoustic reflex measurement. Intelligent Hearing system Evoked potential instrument was used to record Cochlear microphonics (CM) and ABR. A calibrated ILO V6 instrument was used to measure DPOAEs.

A PC with Matlab version (2009) and Adobe Audition version 3 software was used for companding the speech stimuli. The speech stimuli, method used for companding and the procedure used to obtain SIS in different conditions are discussed below.

Test stimuli

In the present study two types of stimuli were used; Sentences and VCV nonsense syllables.

Sentences: Two lists of sentences were taken from quick SIN sentence test in Kannada developed by Methi, Avinash and Kumar (2009). Each list contains 7 sentences and each sentence has 5 key words. The sentences were spoken by a male native speaker of Kannada and was digitally recorded in an acoustically treated room using a unidirectional microphone kept at a distance of 10 cm from the speaker's mouth. Adobe audition software (version 3) was used to record the stimuli. The recorded sentences were normalized so that all the words in a sentence had equal intensity.

Speech shaped noise was used to generate sentences with different SNRs. Speech shaped noise was used as it was made to have the same long term average spectrum as sentences had. It was generated from the whole set of sentences at a sampling frequency of 44.1-kHz by estimating the long-term power spectrum of recorded test sentences. This was done by randomizing the phase of the Fourier spectrum of concatenated words of original signals using MATLAB (The Math Works, Natick, MA, USA) software (version 2009). It had a spectrum which approximates the long term average spectrum of the target sentences spoken by an adult male with a secondary peak presented around 100 Hz. Different SNRs were generated using MATLAB. In each list, first sentence were recorded without noise, second sentence was recorded at +15 dB SNR, third sentence at+10 dB SNR, fourth sentence at +5 dB SNR, fifth sentence at 0 dB SNR, sixth sentence at -5 dB SNR, and last sentence was recorded at -10 dB SNR. This was done as the sentences were used to obtain minimum SNR at which 50% word correctly identified in a sentence. The rms level of all these noises was adjusted according to the level of the target speech. The noise and speech was added prior to companding process.

VCV nonsense syllables: Twenty Vowel-Consonant-Vowel syllables (VCV) comprising of the kannada consonants /k, g, ch, t, d, th, dh, n, p, b, m, j, r, v, s, sh, y, h, l, l./ in the context of the vowel /a/ were used. These syllables were spoken by a female native speaker of Kannada and digitally recorded. The data acquisition system had a sampling frequency of 44.1 kHz and 32 bit analogue-to-digital converter.

The 20 VCV syllables were randomized to form 4 lists. Speech shaped noise was used to generate VCV syllables with different SNRs. List 1 was recorded without any noise. List 2 was recorded at +15 dB SNR. List 3 was recorded at +10 dB SNR. List 4 was recorded at 0 dB SNR.

The intelligibility of these recorded stimuli was established by obtaining a speech intelligibility rating on 10 normal hearing young Kannada speaking adults. Only those recorded stimuli judged as having good intelligibility were considered for the study.

These sentences and VCV syllables were spectrally enhanced using companding both in quiet and at different SNRs. The companding architecture was implemented in MATLAB. These companded test stimuli were named as modified sentences and VCV syllables. The companding was done after the mixing of speech noise and speech stimuli because, if this is found to be beneficial for individuals with AD, then the same can be suggested to be incorporated in amplification devices. The devices with this type of strategy incorporated would process noise and speech together as it would receive both simultaneously.

The strategy used a non-coupled filter bank and compression-expansion blocks as shown in Figure. 1. Every channel in the companding architecture had a relatively broad prefilter, a compression block, a relatively narrowband postfilter, and an expansion block. The prefilter and postfilter in each channel had the same center frequency. The pre and postfilter banks had logarithmically spaced center frequencies that span the desired spectral range.

First, the incoming signal was divided into a number of

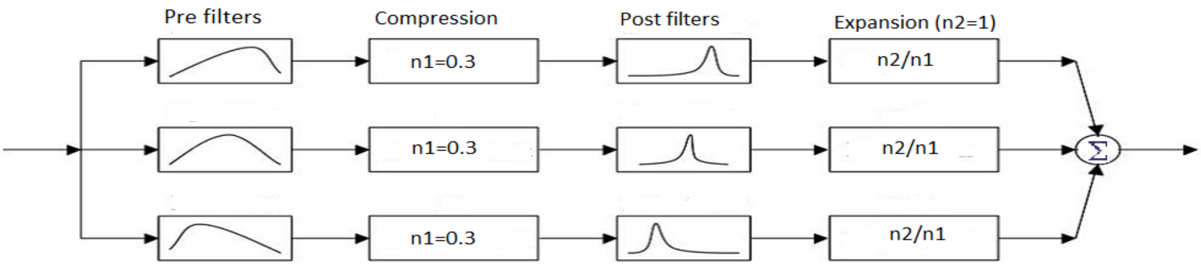


Figure 1: Block diagram of the companding architecture, showing the stimulus analyzed in a bank of broad band prefilters. The output of each prefilter was then subjected to compression, and output was filtered again using sharper postfilters, before it was expanded. The outputs from each channel are then summed to produce the processed broadband companded stimulus.

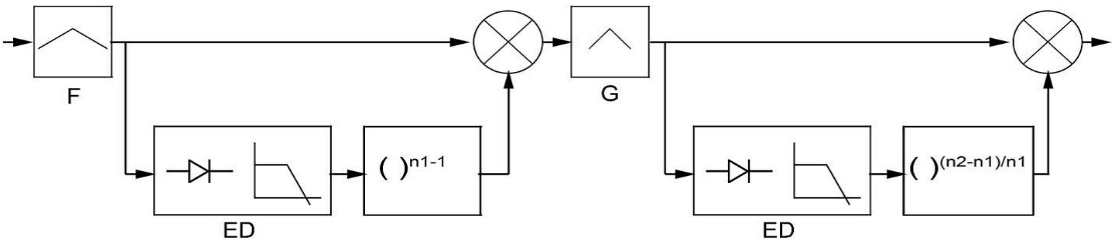


Figure 2: Detailed architecture of a single channel envelope detector.

frequency channels by a bank of relatively broad bandpass filters *F*. Figure. 2 shows the detailed architecture of a single channel companding pathway. The signal within each channel was then subjected to amplitude compression. The extent of compression was dependent on the output of the envelope detector, (ED), and the compression index, (n 1). The compression index $n1$ had a value of 0.3. The compressed signal was then passed through a relatively narrow bandpass filter *G* before being expanded. The gain of the expansion block depends on the corresponding ED output and the ratio of n 2 /$n1$. The $n2$ parameter of the algorithm is expansion index and had a value of 1. The outputs from all the channels were summed to obtain the processed signal. The Adobe audition software (Version 3) was used to normalize the test stimuli to a level of -15dB. After implementing the companding process, there were altogether 8 VCV syllables lists (4 with & 4 without companding VCV syllables list) and 4 sentence test lists (2 with & 2 without companding process). These sentence lists and the VCV syllables lists were transferred digitally to a recordable compact disc. A calibration tone of 1 kHz with a level matched to the normalized level of the stimuli, was recorded prior to each list, using the 1-kHz calibration tone, VU meter on the audiometer was adjusted to read '0'.

Test Procedure

Speech recognition experiments were conducted in normal hearing listeners and individuals with Auditory Dys-synchrony. The test stimuli were played manually by a PC and were routed to a calibrated diagnostic audiometer (GSI-61) which was presented to the participants monaurally at their most comfortable level through the TDH 50 headphones. In the sentence recognition tests and the VCV syllable recognition tests, the participants were presented with a target sentence and the VCV syllables respectively. They were told that, they would first hear the stimulus without any background noise and then with a noise background. They were instructed to repeat or write the stimuli, after the stimuli were heard.

All the participants with normal hearing had to take a pretest consisting of syllables and sentences (with and without companding) at 0dB SNR and participants who scored 60% and above for sentences and 40 % or above for the VCV syllables were taken as participants in the study.

All VCV syllables lists (companded and without companding) at different SNRs were presented and numbers of syllable correctly identified were obtained for both the clinical as well as the control group. Without any modification, VCV syllables were presented in quiet and the number of syllables correctly identified was obtained in both the groups (clinical and control group). Modified (spectrally enhanced through companding) VCV syllables were presented in quiet and number of syllable correctly identified was obtained in both the groups. The procedure was repeated for different SNRs (+15, +10, +0 dB). For the sentence identification, the parameter was the minimum SNR that resulted in more than 50% of words being correctly identified (SNR-50).

Table 2: Mean and standard deviation (SD) of number of VCV syllables correctly identified (syllable identification scores) at 4 different SNRs (quiet, +15, +10, 0 dB), in two stimulus conditions (with and without companding) in the control group.

Stimulus	Mean number of syllables	SD
VCV at quiet without companding (VCVq)	20.00	0.00
VCV at quiet with companding(VCVqC)	20.00	0.00
VCV at +15 dB SNR without companding (VCV15)	19.53	0.74
VCV at +15 dB SNR with companding (VCV15C)	19.13	1.30
VCV at +10 dB SNR without companding (VCV10)	18.20	1.37
VCV at +10 dB SNR with companding (VCV10C)	18.00	1.51
VCV at 0 dB SNR without companding (VCV0)	10.47	2.13
VCV at +15 dB SNR with companding (VCV0C)	13.00	2.13

Out of the 2 companded sentence lists, one list was presented in quiet initially and number of word correctly identified was noted. If they had correctly identified at least 50% words, then the SNR was reduced to + 15 dB and a different set of sentences were presented once again. The same procedure was followed till they achieved the minimum SNR at which they could correctly identify 50% of the words in a sentence and that SNR value was noted. A similar procedure was used for the non companded list. Sentence lists were chosen randomly to minimize the list bias. The noise conditions in all the steps were presented in the order of increasing level of difficulty.

Results

There were basically two parameters of interest: a) the number of syllables correctly identified on presenting VCV syllables (companded and non companded) in quiet and noise in both the groups and b) the minimum SNR at which 50% of the words were correctly identified in a sentence (companded and non companded) in both the groups. Descriptive and inferential statistics were carried out over both these parameters in the control and the clinical groups which were then compared between the two groups.

Control group

A total of 15 ears with normal hearing sensitivity were included. The mean and standard deviation for the number of VCV syllables correctly identified across the two stimulus conditions (with and without companding) at four different SNRs were calculated and the details are given in the Table 2.

It is evident from Table 2 that the syllable identification improved as the SNR improved for both companded and non companded stimuli. The control group performance reached the ceiling in quiet. At +15 dB SNR and +10 dB SNR, non companded stimuli show slightly bet-

Table 3: Bonferroni adjusted multiple comparison test results for syllable identification scores obtained between any two SNRs (quiet, +15, +10, 0 dB) in the control group

(I) SNR	(J) SNR	Mean Difference (I-J)	Sig.
Quiet	+15dB SNR	0.667	0.14
	+10dB SNR	1.900	0.00
	+0dB SNR	8.267	0.00
+15dB SNR	+10dB SNR	1.233	0.01
	+0dB SNR	7.600	0.00
+10dB SNR	+0dB SNR	6.367	0.00

Table 4: t-value, degree of freedom and level of significance for pair wise comparison of syllable identification scores between stimulus conditions (with and without companding) and between different SNRs in the control group

Comparison	t-value	df	Sig
VCVq - VCV15	2.43	14	.029
VCVq - VC10	5.08	14	.000
VCVq - VCV0	17.31	14	.000
VCV15 - VCV10	4.40	14	.001
VCV15 - VCV0	15.18	14	.000
VCV0 - VCV0	12.31	14	.000
VCVqC - VCV15C	2.58	14	.022
VCVqC - VCV10C	5.12	14	.000
VCVqC - VCV0C	11.64	14	.000
VCV15C- VCV10C	3.01	14	.009
VCV15C - VCV0C	8.99	14	.000
VCV10C - VCV0C	8.79	14	.000
VCV15 - VCV15C	2.45	14	.028
VCV10 - VCV10C	.642	14	.531
VCV0 - VCV0C	5.429	14	.000

Table 5: Mean and standard deviation (SD) of minimum SNR at which in minimum 50% of the words correctly identified in sentence test materials, with and without companding in the control group

Conditions	Mean (SNR)	N	SD
Minimum SNR without companding	-2.67	15	2.58
Minimum SNR with companding	-5.00	15	0.00

Table 6: Mean and standard deviation (SD) of number of syllables correctly in the two conditions (with and without companding) in quiet and +15 dB SNR in the clinical group

	VCVq	VCVqC	VCV15	VCV15C
Mean(No. of syl)	10.67	12.47	3.33	4.33
SD	6.275	6.413	4.451	5.912

ter scores while at 0 dB SNR, companded stimuli appear to be better.

To know whether there was any significant difference in performance of the control group at different SNRs and in the two stimulus conditions (with and without companding), two way repeated measure ANOVA was done. A significant main effect on syllable identification was seen at different SNRs [F (3, 42) = 143.72, $p < 0.05$]. A significant main effect was also observed for companded and non companded conditions [F (1, 14) = 8.83, $p < 0.05$]. A significant interaction between SNRs and conditions [F (3, 42) = 24.11, $p < 0.05$] was present. Bonferroni's adjusted multiple comparison test was used as the post-hoc test and the details are shown in Table 3. The Bonferroni adjusted multiple comparison test revealed a significant difference in syllable identification between all SNR conditions except between quiet and +15 dB SNR condition. Paired sample t-test was used to test significant difference in syllable identification between the conditions and also between the SNRs. Details of the paired sample t- test is shown in Table 4. Paired sample t-test results showed indicated a significant difference in all the conditions expect at +10 dB SNR where the control group performed equally well in with and without companding.

The mean and standard deviation of the SNR-50 for sentences with and without companding in the control group are given in Table 5. It is evident from Table 5 that the SNR-50 for sentences was lower (better) for companded condition compared to non companding condition. To test for significance, a paired sample t-test was done. A significant difference was seen [$t(14) = 3.50$, $p < 0.05$] between the two conditions indicating that performance was better with companding than without companding in the control group.

Table 7: Z-value and level of significance for pairwise comparison of syllable identification between two stimulus conditions (with and without companding) and 2 different SNRs (quiet, +15 dB) obtained in the clinical group

	VCVqC -VCVq	VCV15C -VCV15	VCV15 -VCVq	VCV15C -VCVqC
Z-value	-3.09	-1.90	-3.19	-3.18
Sig.	0.00	0.06	0.00	0.00

Clinical group (Auditory Dys-synchrony)

A total of 15 ears (10 individuals) having auditory dyssynchrony comprised the clinical group. The clinical group participants could not identify any of the syllables at +10 dB SNR and 0 dB SNR. The mean and standard deviation for the syllable identification scores across the two stimulus conditions (with and without companding) in quiet and at +15 dB SNR are given in Table 6. The number of syllables identified was more for companded stimuli than the non companded stimuli, both in quiet as well as at +15 dB SNR. Due to a high standard deviation, non parametric tests were used for the clinical group to assess significance. To know whether the performance of the participants in the clinical group differed significantly across the two stimulus conditions and between the two SNRs, Wilcoxon Signed Rank Test was carried out. Table 7 details the test results. As evident from Table 7, Wilcoxon Signed Rank Test showed a significant difference in syllable identification in all the conditions except at +15 dB SNR, where the difference between companded and non companded syllable identification was not significant (p=0.06).

Table 8: Mean and standard deviation (SD) of minimum SNR at which 50% of the word correctly identified in sentence test materials, in two conditions (with and without companding) in the clinical group (n=13)

	Mean(SNR)	SD
Minimum SNR without companding	4.62	2.47
Minimum SNR with companding	0.77	3.44

For the sentence test material, only 13 ears were tested as data from two ears could not be obtained for the task. The mean and the standard deviation for this data are shown in Table 8. The data appear to indicate better performance in the companded condition. To test for significance, Wilcoxon Signed rank test was carried out. A significant difference was found between the two conditions [$Z = -2.89$, $p < 0.05$] indicating that the clini-

Table 9: Z- value and level of significance for comparison of VCV syllable identification scores between the clinical and control groups obtained at 2 SNRs (quiet, +15 dB) and in two conditions (with and without companding)

	VCVq	VCVqC	VCV15	VCV15C
Z- value	-4.99	-4.99	-4.83	-4.83
Sig.	0.000	0.000	0.000	0.000

Table 10: Z- value and level of significance for comparison of minimum SNR that resulted in 50% of the words correctly identified in sentence test materials, in two conditions (with and without companding) between the clinical and the control groups

	Minimum SNR without companding	Minimum SNR with companding
Z- value	-4.40	-4.73
Sig.	.00	.00

cal group indeed performed better with companded sentences when compared to non companded sentences.

Between Group Comparisons

Comparison of number of syllables correctly identified across the clinical and control group at different SNRs in two conditions (with and without companding): Mann Whitney test was used to compare between the two groups across the SNRs and between companded and non companded syllables. The details can be seen in Table 9. The results of Mann-Whitney Test showed significant difference in performance between control and clinical group for syllable identification at both the SNRs and also in between the two conditions (with and without companding). Thus, the control group performed significantly better than clinical group in VCV identification task at all the SNRs in both stimulus conditions.

Comparison of the SNR-50 for sentences between the clinical and the control group with and without companding : Mann Whitney test (Table 10) revealed a significant difference between the two groups with and without companding. The control group, thus performed better at low SNRs than the clinical group in both companded and non companded conditions.

Discussion

Findings in Control group

In the present study, the number of syllables correctly identified significantly reduced as the SNR reduced. The number of syllables identified was the least for the non companded signal at 0 dB SNR and maximum number of syllables correctly identified was seen in the quiet condition.

The results obtained of the present study are similar to those obtained by earlier investigators (Dorman, Loizou & Tu, 1998). Houtgast and Steeneken (1985) also reported that speech intelligibility is reduced in the presence of background noise. This is partly because the noise reduces the modulations of speech envelope. In addition, the decline in intelligibility may also result from the distortion of temporal fine structure and introduction of spurious envelope modulation, as these modulations obscure or mask the modulation pattern of speech, and obliterate some of the cues for identification (Drullman, 1995).

The individuals in the control group demonstrate a significant benefit from companded VCV syllables at 0 dB SNR. This shows that companding enhances the spectral peaks and listeners could take advantage of these enhanced peaks in adverse listening conditions. The result of the present study is in agreement with the results of previous investigators where they showed significant improvement in identification scores in the presence of background noise, when envelope enhanced stimuli were presented to individuals with normal hearing and cochlear hearing loss (Apoux et al., 2004; Baer et al., 1993; Bunnell, 1990; Clarkson & Bahgat, 1991; Franck et al., 1999; Lyzenga et al., 2002). Turicchia and Sarpeshkar (2005) also showed that spectral contrast is an emergent property of the companding strategy and had speculated that this strategy has the potential to improve speech performance in noise.

Studies have found small improvement in the identification of stop consonants in quiet, using a contrast enhancement technique in which the envelope amplitude of each Fast Fourier Transform bin is enhanced in proportion to the difference in the original envelope amplitude and the average spectrum level (Bunnell, 1990; Franck et al., 1999; Lyzenga *et al.*, 2002). Baer *et al.* (1993) convolved the spectrum with a difference of Gaussian filter to provide spectral enhancement. They showed that their normal hearing participants preferred speech in noise with moderate enhancement in terms of quality and intelligibility. In the quiet condition, the performance of normal hearing individuals had already reached the ceiling without companding. So, the benefits of companding could not be seen. In situations with good SNR, companding is not expected to provide additional benefit since the individuals with normal hearing do not have any kind of spectral and temporal deficit. But companding will be effective in adverse listening conditions, which is shown by a better performance at 0 dB SNR in the companded condition.

The results of the normal hearing individuals averaged across 15 listeners showed that for companded sentences the minimum SNR that is required to correctly identify 50% of the words is achieved at significantly lower SNR (better) than without companded sentences. This shows better performance at lower SNR for with companded stimuli compared to without companding sentence test material.

These findings are consistent to earlier investigator's results who reported average improvement in sentence and word recognition in noise, in normal hearing participants using companding strategy (Bhattachrya & Zeng; Oxenham et al., 2007). Normal hearing individuals, in the adverse listening conditions utilized the enhanced spectral and temporal contrast of the companded speech stimuli. The improvement observed in the present study for the companded stimuli can be attributed to the increased spectral and temporal contrast provided by companding strategy.

Findings in Auditory Dys-synchrony

Individuals with AD also showed the trend of reduced number of syllable identification with poorer SNR. Participants in the clinical group could only perform at two SNRs, that is in quiet and at +15 dB SNR. This suggests that listeners with AD have difficulty in utilizing available information if the condition is even slightly worse. Studies have investigated speech perception in noise in individuals with AD and have reported that noise has a detrimental effect on speech perception (Rance et al., 2007; Zeng & Liu, 2006). The results in the present study are in accordance with the previous studies (Rance et al., 2007; Zeng & Liu, 2006). Adding noise to the speech signal leads to problem in perceiving the envelope of speech, because of reduction in modulation depth and addition of spurious modulation (Drullman, 1995). This explanation would explicate severe degradation in speech intelligibility in the presence of background noise for individual with AD.

The exact mechanism causing extreme difficulty in understanding speech in the presence of noise in individuals with AD is unclear. Zeng et al. (2005) reported an excessive masking effect for the detection of tones in the presence of noise. This excessive masking may be one of the factors in these individuals contributing to the extreme difficulty in understanding speech in the presence of noise.

The results of the present study unequivocally demonstrated benefit from companding. Companded VCV syllables significantly improved syllable identification in quiet. Companding increases the spectral and temporal contrast in speech (Bhattacharya & Zeng; 2007, Oxenham et al., 2007). Loizou (2005) implemented companding strategy in CI users and found a modest improvement in vowel recognition. Narne and Vanaja (2008) have shown that enhancing the envelope improved consonant identification at quiet in individuals with AD. In addition, Zeng and Liu (2006) have said that the participants with AD showed improved performance, in quiet when clear speech is presented and this improvement is attributed to enhanced envelopes in the clear speech. Thus, it can be said in the present study that, the enhanced spectral and temporal contrast for the companded stimuli might be the reason for the improved speech perception in individuals with AD.

There was a difference in performance between with companding (mean=4.33) and without companding (mean=3.33) VCV syllables at + 15 dB SNR, but it was not statistically significant. Lack of significant difference at + 15 dB SNR suggests that, individuals with AD cannot utilize enhanced information in the presence of noise which indicates that they have more of neural problem, which predominantly exhibit temporal deficit.

The results of the individuals with AD averaged across 15 ears showed that for companded sentences the minimum SNR that is required to correctly identify 50% of the words is achieved at significantly lower SNR (better) than without companded sentences. Hassan (2011) reported that temporal modification of speech is beneficial for participants with AD. These results are consistent with earlier investigators who reported an average improvement in sentence and word recognition in noise in individuals with AD using enhanced envelope cues (Zeng & Liu, 2006). Companding enhances the spectral peaks and listeners are able to take advantage of these enhanced peaks in adverse listening condition (Bhattachrya & Zeng, 2007; Oxenham, Simonson, Turicchia & Sarpeshkar, 2007). Bhattachrya and Zeng (2007) reported that companding apart from improving spectral contrast also enhances the temporal contrast. As the individual with AD exhibit temporal deficits enhancing the spectral and temporal contrast might have lead to significant improvement.

Findings in Between Group Comparisons

Syllable identical scores in the presence of noise were significantly more affected in individuals with AD when compared to listeners with normal hearing. Studies have reported that normal hearing listeners use fine structure in understanding speech in adverse listening conditions (Zeng & Liu, 2006). However, individuals with AD are impaired in extracting both envelope and fine structure cues from speech signal even in quiet, adding noise to the speech signal may exaggerate their problem in perceiving the envelope of speech, because of reduction in modulation depth and addition of spurious modulation (Drullman, 1995). In the present study, probably the impaired ability to extract the envelope and fine structure cue could be the reason for their poorer performance compared to normal hearing individuals.

Another possible reason for poorer performance in presence of noise in individuals with AD compared to normal hearing individuals might be due to the excessive masking effect in individuals with AD (Zeng et al., 2005). They could not utilize any information in the presence of noise which indicates they have more of neural problem, which predominately exhibits temporal difficulty.

There was improvement in both the group for companded VCV syllables perception, but the performance of the normal group was significantly better compared to clinical group for syllable identification at all the SNRs. The results suggest that individuals with auditory dys-synchrony were unable to fully utilize the temporal and spectral cues. Potential reasons behind this can be attributed to the poor temporal and spectral processing abilities (Rance et al., 2004; Zeng et al., 2005).

Secondly, spectral enhancement provided by the companding was across frequencies, including low frequencies. The low frequencies instead of enhancing speech perception might have caused upward spread of masking, thus the benefit which individuals with AD got from high frequency enhancement also reduced causing minimal improvement in their SIS compared to normal hearing individuals. Improvement in both the group, for the minimum SNR that is required to correctly identify 50% of the words in a sentence for the companded sentence test stimuli compared to without companded stimuli was observed, but the performance of the normal group was significantly better compared to clinical group. Apoux et al. (2004) have clearly shown that envelope enhancement enhances the consonant portion and comprises the vowel portion of the signal and improves the perception of speech in noise better than other signal processing strategies in individuals with normal hearing. In addition, Picheny, Durlach and Braida (1985, 1986) have said that advantage of clear speech over conversational speech in noise for cochlear hearing loss listeners may be due to the increased consonant to vowel ratio and enhanced envelopes. Thus, it can be said that the improvement observed in the present study may probably be due to enhanced spectral and temporal contrast.

A possible reason for the poorer performance of individuals with AD compared to normal hearing individuals is their inability to fully utilize the temporal and spectral cues. In AD, the neural temporal processing is disrupted, which affects the listener's ability to cope with the dynamic nature of speech signal. Severe disruption of timing cues could impair not only the ability to use amplitude envelope cues in speech but also to perceive rapidly changing spectral shapes in the flow of speech stimuli (Rance et al., 2004).

To conclude, speech perception is significantly impaired in individuals with AD. This is probably a reflection of excessive masking and diminished temporal processing abilities. Enhancing the spectral and temporal contrast through companding may improve speech perception in these individuals. Utilizing the companding strategy in hearing aids may provide benefit in individuals with AD.

Conclusion

Our behavioral data suggest that companding the speech signal enhances the spectral and temporal contrast and lead to better speech perception in individuals with auditory dys-synchrony. There are as many studies which have reported that hearing aids have failed to show beneficial effects in participants with auditory dys-strategies improve speech perception in auditory dys-synchrony. The present study provides indirect evidence that hearing aids incorporating companding strategy can enhance spectral and temporal contrast and be beneficial to persons with auditory dys-synchrony. Thus, it is suggested that companding of signal can be used as one of the signal processing strategies improve speech perception in auditory dys-synchrony.

References

American National Standards Institute. (1991). *Maximum Permissible Ambient Noise Levels for Audiometric Tests Rooms,* ANSI S3:1- (1991). New York: American National Standards Institute.

Apoux, F., Tribut, N., Debruille, X., et al. (2004). Identification of envelope expanded sentences in normal-hearing and hearing-impaired listeners. *Hearing Research, 189*, 13-24.

Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times. *Journal of Rehabilitation Research and Development, 30,* 49-72.

Berlin, C. I., Hood, L. J., Hurley, A., & Wen, H. (1996). Hearing aids: only for hearing impaired patients with abnormal otoacoustic emissions. In C.I. Berlin (Ed.), *Hair cells and hearing aids*. (pp. 99-111). San Diego: Singular Publishing groups.

Bhattacharya, A., & Zeng, F. G. (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noise. *Journal of the Acoustical Society of America, 122,* 1079-1089.

Bunnell, H. T. (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *Journal of the Acoustical Society of America, 88*, 2546-2556.

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

Clarkson, P., & Bahgat, F. (1991). Envelope expansion methods for speech enhancement. *Journal of the Acoustical Society of America, 89,* 1378-1382.

Dorman, M., Loizou, P., & Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processor with 6-20 channels. *Journal of the Acoustical Society of America, 104*, 3583-3585.

Drullman, R. (1995). Speech intelligibility in noise: relative contribution of speech elements above and below the noise level. *Journal of the Acoustical Society of America, 98*, 1796-1798.

Franck, B. A. M., van Kreveld-Bos, C. S. G. M., Dreschler, W. A., & Verschuure, H. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *Journal of the Acoustical Society of America, 106,* 1452-1468.

Houtgast, T., & Steeneken, H. T. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibity in auditorium. *Journal of the Acoustical Society of America, 77,* 1069-1077.

Hassan, D. M. (2011). Perception of temporally modified speech in auditory neuropathy. *International Journal of Audiology, 50,* 41-49.

Kraus, N., Bradlow, A. R., Cheatham, M. A., Cunningham, J., King, C.D., & Koch, C.D. (2000). Consequences of neural asynchrony: A case of auditory neuropathy. *Journal of Association for Research in Otolaryngology, 1,* 33-45.

Kruase, J. C., & Braida, L. D. (2004). Acoustical properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America, 115,* 362-378.

Kumar, U. A., & Jayaram, M. (2005). Auditory processing in individuals with auditory neuropathy. *Behavioral and Brain Functions,*1-21. doi:10.1186/1174-9081-1-21.

Loizou, P. (2005). Evaluation of the companding and other strategies for noise reduction in cochlear implants. *Conference on Implantable Auditory Prosthesis*, Asilomar, Monterey, California.

Lyzenga, J., Festen, J. M., & Houtgast, T. (2002). A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity. *Journal of the Acoustical Society of America, 112,* 1145-1157.

Methi, R., Avinash, & Kumar, U. A. (2009). Development of sentence material for Quick Speech in Noise test (Quick SIN) in Kannada. *Journal of Indian speech and Hearing Association, 23*(1), 59-65.

Narne V. K. & Vanaja C. (2008). Speech identification and cortical potentials in individuals with auditory neuropathy. *Behavioral and Brain Function, 31*, 4 - 15.

Oxenham, A. J., Simonson, A. M., Turicchia, L., & Sarpeshkar, R. (2007). Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing. *Journal of the Acoustical Society of America, 121*, 1709-1716.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech Language and Hearing Research, 28*, 96-103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech Language and Hearing Research, 29*, 434-446.

Rance, G., Cone-Wesson, B., Wunderlich, J., & Dowell, R. (2002). Speech perception and cortical event related potentials in children with auditory neuropathy. *Ear and Hearing, 23*, 239-253.

Rance, G., McKay, C., & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing, 25,* 34-46.

Rance, G., Barker, E. J., Mok, M., Dowell, R., Ricon, A., & Garratt, R. (2007). Speech perception in noise for children with auditory neuropathy/ dys-synchrony type hearing loss. *Ear and Hearing, 28,* 351-360.

Rance, G., & Barker, E.J. (2008). Speech perception in children with auditory neuropathy/dyssynchrony managed with either hearing AIDS or cochlear implants. *Otology Neurotology, 29*, 179-182.

Rance, G., Fava,R., Baldock, H., Chong, A., Barker, E., Corben, L, & Delatycki, M. (2008). Speech perception in individuals with Fredeich ataxia. *Brain, 131*, 2002-2012.

Sininger, Y., & Oba, S. (2001). Patients with auditory neuropathy: Who are they and what can they hear? In Y. Sininger, & A. Starr (Eds.), *Auditory neuropathy: A new perspective on hearing disorder*, (pp. 15-36). Canada: Singular publishing group.

Starr, A., Picton, T.W., Sininger, Y., Hood, L., & Berlin, C.I. (1996). Auditory neuropathy. *Brain, 119*, 741-753.

Turicchia, L., & Sarpeshkar, R. (2005). A bio-inspired companding strategy for spectral enhancement. *IEEE Transactions on Speech and Audio Processing, 13*, 243-253.

Zeng, F. G., Oba, S., Garde,S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in Auditory Neuropathy. *NeuroReport, 10*(16), 3429-3435.

Zeng, F. G., & Shannon, R. V. (1999). Psychophysical laws revealed by electric hearing. *NeuroReport, 10*, 193l-1935.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology, 93,* 3050-3063.

Zeng, F. G., & Liu, S. (2006). Speech perception in auditory neuropathy participants. *Journal of Speech and Hearing Research, 42*(2), 367-380.

Subcortical Encoding of Manipuri Pitch Contours inNative Speakers and Non-Native Speakers

[1]Sneha Bansal & [2]Vijaya Kumar Narne

Abstract

The aim of the present study was to examine the subcortical representation of pitch contours of Manipuri language in native Manipuri-speakers and in non-native speakers with and without musical training. The study also aimed to compare the FFRs recorded in native Manipuri speakers, with and without musically trained, non-native speakers. Three Manipuri disyllabic words were spoken by a native speaker and the three tonal contours (rising, falling and flat) were extracted. These were used as stimuli to extract the FFRs in the three groups of subjects. Pitch tracking accuracy and pitch strength were calculated of the response obtained from each of the groups. Overall, the results demonstrated that the pitch strength was significantly better in the Manipuri group, followed by musically trained group and then the non-manipuri group for rising and flat tones.

Keywords: *Frequency Following Response, tonal contours, Pitch tacking accuracy, Pitch strength.*

Introduction

Tonal languages are those in which pitch variations are used to indicate different meanings at the word or syllable level.Such languages are common in the Far East, South-East Asia (i.e Thai, Cantonese, Mandarin and Taiwanese) and West Africa (Yoruba). Languages in which pitch variations are usually not leading to different meanings at the syllable or word level (for example, English, Hindi, and Kannada) are called as non-tonal languages. In these languages pitch variations indicate stress, different intonation patterns at post lexical levels. Tonal languages provide an optimal window for investigating how long-term experience with time-varying pitch patterns shape perceptual and neural processing of pitch.

Earlier investigators believed that the language processing and plasticity are confined to the cortical structures but they are not a part of the brainstem encoding (Hickok &Poeppel, 2004; Zatorre, Evans, Meyer, &Gjedde, 1992).Speech specific operations may not begin until the signal reaches the cerebral cortex (Scott &Johnsrude, 2003).

However, Galbraith, Arabagey, Branski, Commerci and Rector(1995), Galbraith et al. (2004), showed increased amplitude of Frequency Following Response (FFR) to forward speech when compared to reversed speech.

These results indicate the familiar phonetic and prosodic properties of forward speech that selectively activates the brainstem neurons. It is hence now being accepted that subcortical structures may contribute actively to auditory processing, and are not simply passive relay stations sending information from the peripheral sensory organs to the cortex. In addition, FFR studies have demonstrated that even at the early subcortical stages of processing, the auditory system is found to be malleable due to interactions between the sensory and cognitive processes (Kraus &Banai, 2007). Thus, over the years, evidence is mounting that the experience dependent neural plasticity is not only limited to the auditory cortex but is also seen in the subcortical structures (Griffiths, Uppenkamp, Johnsrude, Josephs & Patterson, 2001).

Krishnan, Xu, Gandour and Cariani (2005) measured the impact of long-term language experience on pitch coding using frequency following response (FFR). They found that native Mandarin-speaking subjects, with an exposure of about twenty years showed more precise linguistic pitch pattern encoding relative to the native English-speaking subjects. Speakers of tonal (Mandarin) language have improved pitch encoding at the level of brainstem compared to non-tonal (English) language speakers irrespective of whether it is a speech or non-speech context (Krishnan, Swaminathan&Gandour, 2009; Swaminathan, Krishnan &Gandour, 2008).

It is well documented that neural encoding of pitch in the auditory brainstem is shaped by long-term experience with language and music (acoustic experience). FFRs between native Mandarin speakers and English speakers have shown that long-term experience with lexical tones improves periodicity encoding at the brainstem (Krishnan, Swaminathan&Gandour, 2009; Krishnan, Xu, Gandour&Cariani, 2005; Swaminathan, Krishnan &Gandour, 2008). These investigators have proposed that the precise pitch pattern encoding is specific to language experience but not an acoustic experience. However, few other investigators, Wong, Skoe, Russo, Dees and Kraus (2007) found that English (non-tonal language) speaking musicians who did not speak Mandarin (tonal-language) showed more robust and faithful encoding of Mandarin tones. The aforementioned studies indicate that both language and acoustical experience (music) would result in a more robust encoding of

[1]Email: snehsmiles@gmail.com,
[2]Lecturer in Audiology, Email: vijaynarne@gmail.com

signal at the brainstem.

According to the studies, tonal languages and music arguably provide the best window for studying how long-term experience with time-varying pitch patterns shapes the neural representation of periodicity. There are only a handful of studies which have examined the extent to which experience dependent effect is specific to a particular language or a general acoustical experience. Also, there are few studies investigating subcortical processing in the tonal languages of India. Therefore, there is a need to examine the long term effects of language exposure and music exposure in the Indian context.

In the present study, Dynamic Iterated Rippled Noise (DyIRN) has been used as the stimuli which allow us to study the neural mechanisms underlying the representation of pitch patterns of those that occur in natural speech without a semantic confound. In the present study, FFRs of native and non-native (with no exposure to music) speakers of Manipuri were recorded and compared to assess the effect of long-term experience with language. The same was carried out with FFRs recorded from musically trained non-native speakers with native speakers of Manipuri, who have not been trained in music to assess if musical experience confers similar benefits as tonal language exposure.

Method

Participants

In the present study, Frequency Following Response (FFR) was recorded from three group of subjects. These groups were native Manipuri speakers as Group I, musically trained non-Manipuri speakers as Group II, and non-musically trained non-Manipuri speakers. All of them had pure tone thresholds within 15 dB HL at octave frequencies between 250 Hz and 8000 Hz. A normal middle ear function was ensured with type-A tympanogram and the presence of bilateral acoustic reflexes. Also, a detailed case history ensured that none of them had any history of middle ear pathology, and did not have complaint of any neurological problem.

Group-I consisted of 10 individuals (5 males & 5 females), who were native speakers of Manipuri, a tonal language spoken in a north-eastern state of India were included in group I. All the subjects in this group were born and raised in Manipur. All of them were in the age range of 18-25 years with mean age 22 years, 3 months.

Group-II consisted of 10 individuals (5 males & 5 females), who were speakers of a non-tonal language were included in group II. None of the subjects in this group were exposed to any tonal language and had no formal musical training. All of them were in the age range of 18-25 years with mean age 19.2 years.

Group-III consisted of 10 participants (3 males & 7 females), who were speakers of a non-tonal language were included in group III. None of the subjects in this group were exposed to any tonal language. They had undergone formal music training for a minimum of five years. All of them were in the age range of 18-25 years with the mean age of 21.1 years.

Test Stimuli

Three disyllabic words of Manipuri were used for recording FFRs. All of the three words were phonetically similar but were distinguished by their tonal contours having three different intonation patterns, rising (Tc1), flat (Tc2), falling (Tc3), signaling three different meanings. The three tonal variations of the stimuli were /tʃaba/-1, /tʃaba/-2 and /tʃaba/-3 which in Manipuri means, "swimming", "eating" and "fitting" respectively. As the tonal variations are lexical for Manipuri-speakers and not for non-Manipuri-speakers, these words were used as the stimuli to test the objectives of the study.

These three words were spoken by an adult female native Manipuri-speaker. They were recorded using a directional microphone in Praat software (version 5.1.31) at a sampling frequency of 44.1 kHz and 16 bit analog to digital convertor. All the stimuli were recorded in a sound treated room.

The three different pitch patterns were extracted using short-term autocorrelation in MATLAB 7. The pitch contours of the three different tones are represented in Figure 1. A polynomial fitting function was employed to fit pitch contour and the equation generated was used in generation of dynamic iterated ripple noise (DyIRN).

Each of these DyIRNs had duration of 250 ms, which included a 10ms cosine ramp to eliminate the spectral

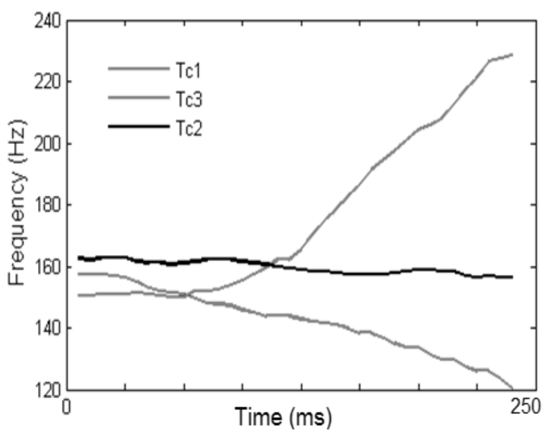


Figure 1: Frequency variations as a function of time of the three different tonal contours, (i.e. Tc1 (Rising), Tc2 (Flat) and Tc3 (Falling)) of the stimulus used in the study.

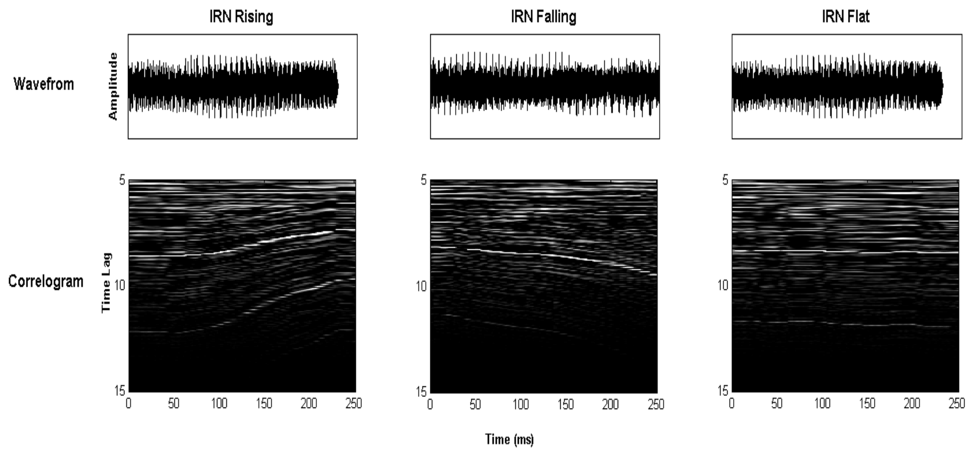


Figure 2: Stimulus waveforms (row 1) and correlograms (row 2), in order from top to bottom, of the IRN homolog of the three different Manipuri tonal contours (rising, falling and flat) of the stimulus used in the study.

splatter and to minimize the onset responses. All the time-varying IRN stimuli were created at a high iteration step (n=12), with a gain of 0.9. Figure 2 shows the waveform, spectrogram and correlogram of the stimuli used in the present study. The generation of IRN noise is provided in the Appendix 1. The stimuli were root mean square normalized to maintain uniform amplitude across all the three stimuli, using group normalization in Adobe audition software (version 3.0). Each stimulus was then converted to a STM file, using Intelligent Hearing System (IHS) stimulus conversion software.

Recording of FFR

Participants were made to relax (preferably sleep) on a reclining chair and they were asked to refrain from any extraneous body movements to minimize movement artifacts. FFRs were recorded from a single channel using gold plated electrodes placed at Cz and referenced to the tip of the nose. An electrode at contralateral mastoid served as ground. The sites of electrode placement were prepared with skin preparation gel and the electrodes were held in their respective positions with a plaster. The absolute electrode impedances were maintained below 5 kΩ and the inter-electrode impedances were less than 2 kΩ in order to maintain low level of noise. The human (FFR) was recorded using the parameters listed in Table 1.

Table 1: Stimulus and acquisition parameters for recording of FFT

Parameters	Stimulus
Duration of stimuli	250 ms
Intensity	70 dBSPL
Transducer	Eartone-3A-insert earphones
Repetition rate	2.76 per second
Number of samples	2000
Filter setting	100 - 3000 Hz
Analysis window	300ms
Mode of presentation	Monoaural
Gain	2,00,000

The order of the stimuli presentation was randomized across the participants. The stimulus files were routed through a digital to analog module and then presented monaurally to the test ear through the magnetically shielded insert earphone (IHS, ER-3A). Each stimulus was presented twice and the stimulus locked responses were subsequently averaged for each stimulus condition to yield a response with a higher signal-to noise-ratio.

Data Analysis

The data obtained was analyzed for the pitch tracking accuracy and the pitch strength of the Manipuri tonal contours.

The ability of the FFR to follow the pitch changes in the stimuli was evaluated by extracting the fo contour from the FFRs using a periodicity detection short term autocorrelation algorithm (Boersma, 1993). Essentially, the algorithm works by sliding a 40 ms window in 10 ms increments over the time course of the FFR. The autocorrelation function will be computed for each 40ms frame and the time- lag corresponding to the maximum autocorrelation value within each frame will be recorded. The reciprocal of this time-lag (or pitch period) represents an estimate of fo. The time- lags associated with autocorrelation peaks for each frame were concatenated together to give a running fo contour. This analysis will be performed on both the FFRs and their corresponding stimuli. Pitch tracking accuracy is computed as the cross- correlation coefficient between the fo contour extracted from the FFRs and the fo contour extracted from the stimuli.

To compute the pitch strength of the FFRs to time varying stimuli, FFRswere divided into six non overlapping 40ms sections (5-45ms; 45-85; 85- 125; 125- 165; 165-

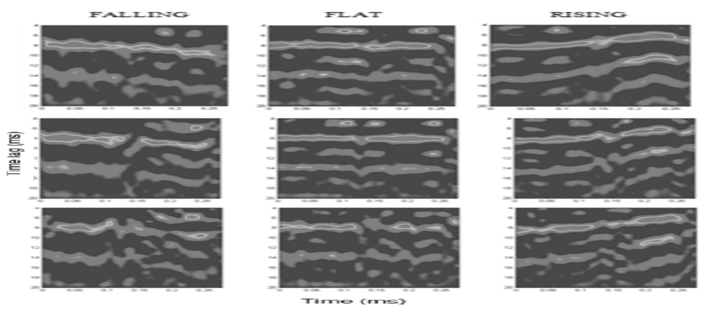


Figure 3: Correlograms derived from the grand average FFR waveforms for the Manipuri group (panel 1), musically trained group (panel 2) and non-tonal group (panel 3) for falling followed by flat and rising tone.

205;205-245ms). The normalized autocorrelation function (expressed as a value between 0 and 1) was be computed for each of these sections, from an analysis of corresponding time frames of the three stimuli and their FFR responses. Here " 0" represented an absence of periodicity and " 1" represented maximal periodicity. Within each 40ms section, a response peak was selected which would correspond to the same location (time-lag) of the autocorrelation peak in the input stimulus (Krishnan et al., 2009; Krishnan et al., 2009; Swaminathan et al., 2008). The magnitude of this response peak represented an estimate of the pitch strength per section. Pitch strength was measured by the average magnitude of the normalized autocorrelation peak per language in each 40 ms frame. All the data analyses were performed using custom routines coded in MATLAB 7 (The Math Works, Inc., Natick, MA). Appropriate statistical analyses were done using SPSS 20.

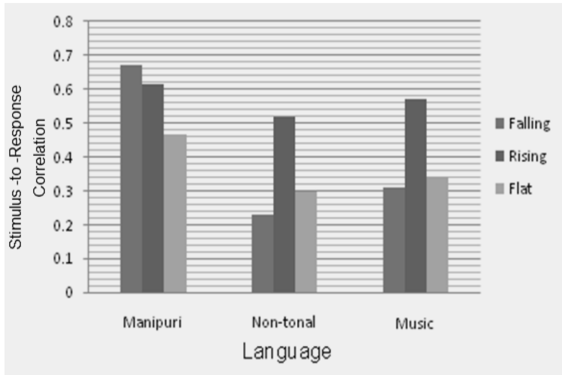


Figure 4: Mean pitch tracking accuracy of FFR for three Manipuri tonal contours (i.e falling, rising and flat) across the three groups of subjects (i.e Manipuri speakers, non-tonal language speakers and musically trained non-tonal speakers).

Results

Temporal and Spectral Properties

Correlograms for grand averaged FFRs are shown in Figure 3 for Manipuri group (panel 1), musically trained group (panel 2) and non-tonal group (panel 3) for falling followed by flat and rising tone. As per the Figure 3, the Manipuri group shows clear dark bands of phase-locked activity at f0 and its multiples in response to the Manipuri tonal contours. The non-tonal group shows that the bands are less distinct and more diffuse across all the tones, while musically trained group showed clearer and darker bands than non-tonal group but lesser than the Manipuri group.

Pitch tracking accuracy of Manipuri tones

FFR pitch tracking accuracy was measured by correlating the response contours in comparison to the stimulus contours for all the three groups. The mean pitch tracking accuracy for three Manipuri tones across the groups were plotted in the Figure 4. As per the Figure 4, pitch tracking accuracy was higher for the Manipuri speaking group as compared to the non-tonal language group and musically trained group.

A Mixed ANOVA on pitch tracking accuracy was performed to assess significant effects of the three Manipuri tonal contours across the three groups of subjects. The analysis revealed, significant main effects of tone [$F(2, 42) = 14.7$, $p<0.001$] and group [$F(2, 21) = 12.7$, $p<0.001$]. The group x tone interaction effects were not significant [$F(2, 21) = 2.4$, $p = 0.11$]. A separate one way ANOVA was carried out to see if the mean difference reaches significance for each of the three tones.

Pitch strength of Manipuri tones

FFR pitch strength was measured by the average magnitude of the normalized autocorrelation peak per group for six tonal sections (5-45, 45-85, 85-125, 125-165, 165-205, 205-245 ms) of 40 ms each, for all the three Manipuri tones and were analyzed. The mean pitch strength for all the groups and each of the three tones i.e. falling, rising and flat are depicted in the Figure 5. As per the Figure 5, the overall pitch strength was found to be greater for the Manipuri group as compared to the non-tonal group.

It can be noted from the Figure 5, for the falling tone, the pitch strength was higher across all the sections for the Manipuri group than the other two groups except at section 5 to 45 ms. For the rising tone, the pitch strength of the Manipuri group and musically trained group was greater than the non-tonal group at all sections except 5 to 45 ms, 45 to 85 ms and 85 to 125 ms. At section 205 to 245 ms, the musically trained group showed lower pitch strength than the Manipuri group. For the flat tone, there was minimal difference in the pitch strength for all the groups except at sections 165 to 205 ms and 205 to 245 ms whereas the Manipuri group and musically trained group showed a better pitch strength than the non- tonal group.

A mixed ANOVA was performed to assess the significant effects of the three Manipuri tones across the three groups of subjects. For falling tone, it revealed significant main effects of group [F (2, 27) = 3.33, p <0.01], sections [F (5, 27) = 5.92, p <0.01] but the group x sections interaction [F (10, 27) =0.97, p = 0.47] was not significant. A one way ANOVA was performed to assess whether the mean difference in pitch strength reaches significance across the groups at each section separately.

The analysis revealed a significant difference across groups for all the sections except at sections 5to 45 ms and 85 to 125 ms. Bonferroni's post-hoc analysis revealed a significant mean difference in pitch strength between Manipuri group and non-tonal group at all the sections except at 5to 45 ms and 85 to 125 ms.

For rising tone mixed ANOVA revealed, no significant main effects of group [F (2,27) =1.44, p = 0.25] and sections [F (5,27) = 1.91, p = 0.09], but the interaction group x section was significant [F (10,27) = 2.32, p <0.01]. Tosee the interaction effect, a one way ANOVA was performed, it revealed a significant difference at section 165 to 205 ms between Manipuri group and non-tonal group (p<0.05). For flat tone mixed ANOVA revealed, no significant main effects of group [F (2,27) = 0.55, p = 0.58], sections [F (5,27) = 0.17, p = 0.31] and group x tone interaction [F (10, 27) = 0.34, p = 0.21]. As there was no main effect of group, sections and their interaction, no further analysis was done. Figure 6. provides a better visualization of the one way ANOVA results, showings at which sections there is a significant difference across the groups for all three Manipuri tones.

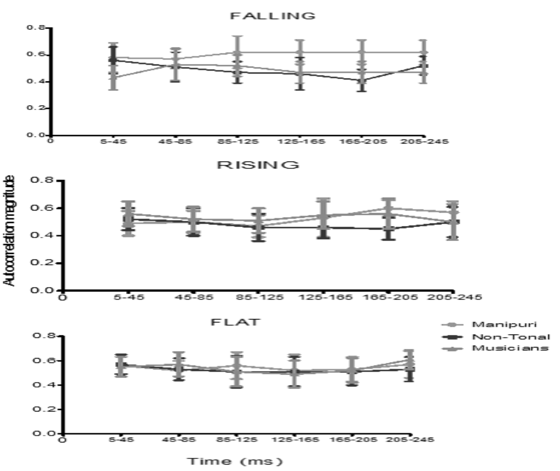


Figure 5: Mean pitch strength of the Manipuri tones within each of their six 40ms sections falling (Panel 1), rising (Panel 2) and flat (Panel 3) as derived from the averaged FFR waveforms of the three groups (Manipuri, Non-tonal language, musically trained).

Figure 6:*The results of one way ANOVA for the three Manipuri tones, falling (column 1), flat (column 2) and rising (column 3) comparing the mean pitch strength between Manipuri group and non-tonal group (row 1) and Manipuri group and musically trained group (row2). The darkened sections show a significant difference amongst the compared groups.*

The results of the present study showed that the pitch tracking accuracy for falling tone was statistically higher in the Manipuri group and musically trained group than the non-tonal group. Of these, Manipuri group showed most accurate pitch tracking compared to other groups. No statistical difference in the pitch tracking accuracy was found across all the groups for the rising and the flat tone. On whole, the pitch strength was more robust for the Manipuri group than the other groups. There was no statistical significant difference across all the subjects for any combination of groups.

For the falling tone, there was significant difference between Manipuri group and non-tonal group at sections 45to 85 ms, 125 to 165 ms, 165 to 205 ms, 205 to 245 ms, but between Manipuri group and musically trained group, only at sections 165 to 205 ms and 205 to 245 ms. Whereas for the rising tone, a significant difference was seen across all the groups only at section 165 to 205 ms. For the flat tone no significant difference was seen in any section across the groups.

Discussion

To accomplish the aim of the present study, brainstem response were obtained using FFRs for three DyIRN, each representing three different tonal contours falling, rising and flat. The results demonstrate that the Manipuri group had greater pitch tracking accuracy and pitch strength than the non-tonal groups. It is further noted that, the falling tone and the rising tone showed more difference for the subcortical pitch representation between tonal group and non-tonal group, whereas for the flat tone, the difference was minimal. Similar to the present study, Krishnan, Gandour, Bidelman, &Swaminathan, (2009) conducted a study comparing the subcortical pitch representation by obtaining FFRs using DyIRN stimuli in tonal (Mandarin) and non-tonal (English) language groups. They also observed similar results as in the present study. However, amount of pitch

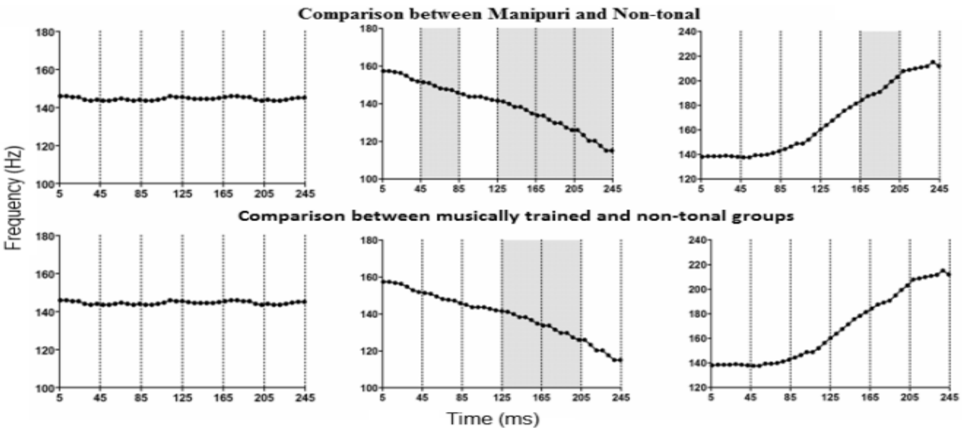


Figure 6: The results of one way ANOVA for the three Manipuri tones, falling (column 1), flat (column 2) and rising (column 3) comparing the mean pitch strength between Manipuri group and non-tonal group (row 1) and Manipuri group and musically trained group (row2). The darkened sections show a significant difference amongst the compared groups.

strength noted in their study was approximately 0.8, whereas in the present study it was 0.6 to 0.7. The precise reasons for reduced pitch strength in the present study are not known. The probable reasons could be the methodological differences, and the instrument used for recording responses. One important difference is the recording system used. The current study used Intelligent Hearing Systems (IHS) Smart EP (Version 4.00) which records at a sampling rate of 16000/sec and averaged waveform has a total of 512 samples. Krishnan et al., (2009) used Tucker-Davis technologies system at a sampling rate of 25000/s and gives 6520 sample for 250 ms.As the number of samples were lesser in the present study, it could have led to a less accurate autocorrelation coefficient value.

In addition, several studies conducted to obtain FFRs using natural or synthetic speech stimuli, to study the subcortical representation of pitch in tonal and non-tonal language speakers, demonstrated similar findings as in the present study (Krishnanet al., 2005; Krishnan et al., 2010).

The results of the present study and previous studies continuously show that linguistically relevant pitch patterns are processed at the brainstem level. The listener's native language experience has changed the way they process linguistically relevant pitch patterns regardless of the stimulus context (non-speech) in which these patterns are embedded. Although the basis for cross-language differences in FFR pitch extraction may emerge from language experience (Manipuri), the effects of such experience are not specific to speech perception (Xu et al., 2006), nor is pitch extraction at the brainstem level necessarily specific to the domain in which pitch patterns occur.

Further it was found that the pitch tracking accuracy in the musically trained group and Manipuri group was not statistically different for the falling tone and the pitch strength for the rising and flat tone. Wong et al., (2007) compared the subcortical pitch representation in terms of pitch tracking accuracy and pitch strength between musicians and non-musicians who were non-native speakers of any tonal language. It was found that the musicians exhibit more faithful pitch tracking and robust encoding of the Mandarin tones than non musicians. This was majorly seen for the falling tone followed by the rising tone. In support to this, the studies conducted to compare the pitch contours with cortical auditory evoked potentials also show similar results (Chandrasekaran, Krishnan, &Gandour, 2009; Fujioka et al., 2004). Further, Song, Skoe, Wong and Kraus (2008) studied the effect of short term linguistic training of pitch tracking accuracy. They noted that there was a decreased number of pitch tracking errors and these decrease was more for the falling contour than the other tonal contours. It can be therefore noted that, along with language experience the long term musical experience also modulates the brainstem processing of musically relevant pitch contours (Munte, Altenmuller, & Jancke, 2002).

The results of the present study and those of previous studies noted that subjects with long term language or musical training have good ability in pitch tracking. Therefore it can be hypothesized that the subcortical pitch processing is shaped by long-term acoustical experience rather than just language experience. That is experience dependent enhancement of pitch representation at the brainstem level is specific to pitch patterns that occur in the listener's acoustical experience. The hypothesis is further supported by studies conducted on comparing two different tonal language groups (Mandarin and Thai) and a non tonal language group (English) using FFR, which reported higher pitch tracking abilities and more robust pitch strength in both of the tonal languages regardless of their language identity, while poorer in non-tonal group (Krishnan, Gandour, &Bidelman, 2010). This implies that the brainstem neurons are differentially sensitive to changes in pitch irrespective of the language identity as long as the two languages have comparable phonological system. Hence the brainstem demonstrates plasticity and the pitch extraction brainstem neurons can be shaped by the acoustical experience of an individual.

Conclusions

From the present study it can be concluded that, FFR can serve as an index of studying subcortical representation of pitch in humans. It can be inferred from the findings that the neural mechanisms underlying the brainstem are adaptive and driven by long-term experience. Over a period of time, they sharpen the pitch extraction neurons for processing pitch contours that are relevant to their language or musical experience. Thus, the brainstem plasticity and pitch representation is the outcome of leaning experience and is not specific to any domain. The study threw light on the nature of subcortical pitch processing in native, non-native and musically trained subjects and the experience dependent plasticity which has implications for future research and understanding the nature of pitch encoding in the human auditory system. . As the results obtained in the present study are in line with most of the research done on subcortical representation of the pitch as reflected by pitch tracking accuracy and pitch strength. It indicates that the both long term and short term learning experience in context of both music and language and is not restricted to either domain.

References

Chandrasekaran, B., Krishnan, A., &Gandour, J. T. (2009b). Sensory Processing of Linguistic Pitch as Reflected by the Mismatch Negativity. *Ear & Hearing*, *30* (5), 552-558.

Galbraith, G. C., Amaya, E. M., de Rivera, J. M., Donan, N. M., Duong, M. T., Hsu, J. N., et al. (2004). Brain stem evoked response to forward and reversed speech in humans. *Neuroreport, 15*(13), 2057-2060.

Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., Rector, P.M. (1995). Intelligible speech encoded in the human brain stem frequency following response. *Neuroreport, 6,* 2363-2367.

Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *NatureNeuroscience, 4*(6), 633-637.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams :a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(12), 67-99.

Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Direction in Psychological Science, 16*(2), 105-110.

Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience, 21*(6), 1092-1105.

Krishnan, A., Xu, Y., Gandour, J., &Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25*, 161-168.

Krishnan, A., Gandour, J., & Bidelman, G.M. (2010). Effect of tone language on pitch processing in the Brainstem. *Journal of Neurolinguistics, 23*, 81-95.

Munte, T. F.,Altenmuller, E., & Jancke, L.(2002).The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience, 3*,473-478.

Song, J.H,.Â Skoe, E., Wong, P.C, & Kraus, N.(2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of cognitive neuroscience. 20*, 1892-1902.

Scott, S. K., & Johnsrude, I. S.,(2003). The neuroanatomical and functional organization of speech perception. *Trends in neurosciences, 26*(2), 100-107.

Swaminathan, J., Krishnan, A., &Gandour, J. T. (2008). Pitch encoding in speech and non-speech contexts in the human auditory brainstem. *Neuroreport, 19*(11), 1163-1167.

Wong PCM, Skoe E, Russo NM, Dees T, & Kraus N. (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10,* 420-422.

Xu, Y., Gandour, J. T., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America, 120*(2), 1063-1074.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*, 846-849.

Acoustic Change Complex: Neural Correlate of Speech in Noise Perception

[1]Spoorthi T. & [2]Devi N.

Abstract

This study investigated the potential of an electrophysiological measure, Acoustic Change Complex (ACC) in detecting the deficits in neural encoding of speech in the presence of noise and it also aimed to identify the neural factors (latencies and amplitudes) that determine good or poor speech perception in noise. Thirty typically developing children, between 12 to15 years participated in the study. They were divided into two groups based on behavioural speech in noise (SIN) scores as, group I (good scores) and group II (poor scores). Latency measures and peak to peak amplitude measures were obtained from ACC recordings for stimulus CV syllable /si/ in quiet and noise conditions of stimulus presentation. Results revealed significant differences between groups irrespective of the conditions and also between conditions irrespective of the groups. A significant correlation was observed between behavioural measure (SIN scores) and electrophysiological measures (latencies & amplitudes). These findings show the feasibility of using cortical responses in understanding individual's difficulty in perceiving speech in noisy background. Both latency and amplitude measures were found to depict speech perception capabilities.

Keywords: *Speech perception in noise, Acoustic change complex.*

Introduction

Speech perception involves interpretation of speech sounds ranging from simple phonemes to complex sentences (Boothroyd, 1997). Speech perception develops in childhood and is influenced by sensory factors, cognitive skills and linguistic abilities (Hnath-Chisolm, Laipply & Boothroyd, 1998). Amongst sensory factors, auditory modality has a major role to play. In other words, for understanding speech, maximum information is obtained through audition. Normal auditory processing of speech is particularly important in children. It provides a solid foundation for acquiring speech and language and in turn, academic skills such as reading and written language (Cunningham, Nicol, Zecker & Kraus, 2000). However, occurrence of auditory perceptual deficits which can impede normal speech perception is not uncommon in them. Such deficits can have adverse effects on language acquisition and in turn, literacy development. Elliott and Hammer (1988) suggested that in some children the root cause for the learning problems is auditory perceptual deficits specifically related to the processing of speech. Treiman, Broderick, Tincoff and Rodriguez (1998) have reported that children, who have auditory problems like difficulty identifying or discriminating phonemes, develop poor spelling and reading abilities.

Some of the auditory perceptual areas where children show deficits include - speech sound discrimination, temporal pattern recognition, auditory integration, localization, lateralization, speech in noise perception. Amongst these, most commonly reported auditory deficit in childhood is difficulty perceiving speech in the presence of noise (Cunningham et al., 2000). This specific perceptual difficulty also has been linked to learning problems. Chermak, Vonhof, and Bendel (1989) found that individuals with learning difficulties have poorer word identification in noise.

Speech consists of dynamic elements that require fine grained neural representation of temporal information. Noise disrupts the neural synchrony required for clear representation of those aspects of speech. This degraded representation of speech in the presence of noise at cortical and sub-cortical levels results in perceptual difficulties (Anderson, Skoe, Chandrasekaran & Kraus, 2010).

These reports from the literature, signifying the deleterious effects of background noise on processing speech, highlights the necessity of abundant research on identifying such problems using behavioral and electrophysiological measures to uncover and confirm those adverse effects. This would help in planning the management options at the earliest. There are a number of subjective tests for identifying poor speech in noise perception in children. Examples - Hearing in noise test (Nilsson, Soli & Sullivan, 1994), Speech Perception in Noise test (Bilger, Nuetzel, Rabinowitz & Rzeczkowski, 1984). But due to factors like cognitive dysfunctions, linguistic limitations, behavioral problems and others, conducting these behavioral tests in children to assess such perceptual deficits may become impractical most of the times. Thus, there is a clinical need to investigate on objective tests which have potential in identifying difficulties related to understanding speech in noise.

Auditory Evoked Potentials have been reported as being potential in reflecting difficulties related to encoding speech in noise. They provide a non-behavioral means

[1]Email: t.spoorthi@yahoo.com
[2]Lecturer in Audiology, Email: deviaiish@gmail.com

of investigating the processing of speech (Ostroff, Martin & Boothroyd, 1998). Utility of sub cortical auditory evoked responses in identifying such deficits have been extensively researched. However, only a few investigations have been conducted on cortical event related potentials. As a result, very little is known about the relationship between central processes and the speech perception in noise. Among the cortical potentials, such studies using ACC are sparse. Thus, this study is an attempt to begin such an investigation to understand the neural encoding of speech in the presence of noise using ACC.

ACC is a P1-N1-P2 complex, elicited by acoustic changes in an ongoing stimulus. Both speech and non speech stimuli can be used to elicit this response. In non speech stimuli - intensity and/or frequency changes or modulations in sustained tones have been reported to elicit ACC (Spoor, Timmer & Odenthal, 1969; Jerger & Jerger, 1970). Also, Ross, Tremblay and Picton (2007) have reported occurrence of this response to inter-aural phase changes in non speech stimuli. In speech stimuli like simple syllables, the transition from consonant to vowel has been shown to elicit this response (Ostroff et al., 1998).

The literature on ACC using speech stimuli has suggested that it provides understanding about auditory system's ability to represent acoustic features present in the speech signal. In support of this, Martin and Boothroyd (2000) have reported that, ACC response can be recorded to the formant frequency changes within a vowel. Further, Martin (2007) reported that this cortical response has good agreement with behavioral frequency discrimination thresholds ($\approx$10Hz). However, whether the objective measures of ACC response can also represent degraded perception of speech stimuli in challenging listening situations like speech in the presence of noise and whether that representation has agreement with behavioral measures has not been investigated.

Method

Participants

Thirty school going children between 12 to 15 years were included. They were divided into 2 groups based on their Speech in Noise (SIN) scores. Group I consisted of 15 children (30 ears) with good speech in noise scores (SIN score $\geq$60%) while Group II included 15 children (30 ears) with poor speech in noise scores (SIN score $\leq$40%). In both the groups, there were 5 children each, in the age range 12 to 13 years, 13 to 14 years and 14 to 15 years. This criterion was to control the maturational effects. Group I consisted of 5 males and 10 females, while group II consisted of 7 males and 8 females. Gender match could not be obtained due to unavailability of the participants.

Children were recruited after obtaining written consent from their parents or guardians. Participants were native Kannada speakers with no history of any neurological, psychological, cognitive or otological problems and normal speech and language development. Air conduction thresholds (at the octave frequencies from 250 Hz to 8000 Hz) and bone conduction thresholds (at the octave frequencies from 250 Hz to 4000 Hz) were $\leq$15 dB HL. Also participants had bilateral normal middle ear function (Type 'A' tympanogram at 226 Hz probe tone and normal ipsilateral & contralateral reflexes at 500 Hz, 1000 Hz, 2000 Hz & 4000 Hz). Speech Recognition Thresholds (SRT) were $\pm$12dB to pure tone average and Speech identification scores in quiet were $>$ 90% at 40 dB SL (ref. SRT). Further, TEOAEs (non-linear clicks of 260 sweeps at 80 dB pe SPL) were present (6 dB SNR & 90% reproducibility) and auditory brainstem response (wave V latency) for click were normal (Repetition rates = 11.1/s & 90.1/s, Intensity = 90 dB nHL). Children did not have any illness on the day of testing.

Test Environment

All testing were carried out in an electrically shielded and sound treated room where noise levels were maintained within permissible limits - ANSI S3.1 (1999).

Test Procedure

Preliminary evaluations: Detailed history regarding otological, neurological, psychological, and cognitive problems was taken along with the details of speech and language development. Once the possible deficits were ruled out in all these areas, pure tone audiometry, speech audiometry, immittance evaluation and TEOAE measurements were carried out. Further, only those children who passed the criteria in all the above evaluations were subjected to Speech in Noise testing.

Speech in Noise (SIN) testing: Phonemically balanced Kannada word lists developed by Vandana (1998) were used. Two lists out of four were considered. Words were presented through monitored live voice at 40 dB SL (ref. SRT) and 0 dB SNR (Speech noise). Twenty five bi-syllables in every list were presented for each trial and every word was given a score of 4%. Children had to repeat the words heard. Number of correctly identified words was noted down to find the SIN score. Children who fell into any of the two groups - Group I (SIN score $\geq$60%) or Group II (SIN score $\geq$40%) were considered for ABR and ACC recordings. This criterion was considered to have good distinction.

ABR and ACC recording: Participants were made to sit comfortably on a reclining chair. They were instructed to sit relaxed without much body and eye movements. They were allowed to watch DVD movies played without sound. Corrosion free silver chloride disc electrodes were used for recording. Absolute impedances were

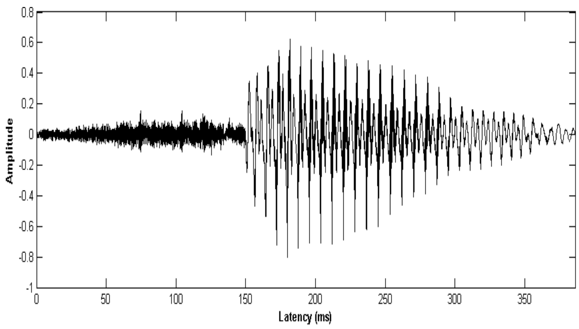


Figure 1: The acoustic waveform of CV syllable /si/ used to record ACC.

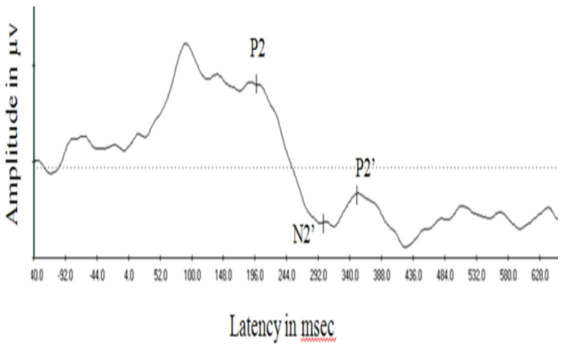


Figure 2: ACC response for syllable /si/ recorded from a typically developing child of age 13 years.

Table 1: Protocol used for ABR and ACC recording

	ABR	ACC
Transducer	ER-3A insert earphones	
Stimulus type	Click	CV Syllable-/si/
Stimulus duration	100 μs	386 ms
Stimulus intensity	90 dB nHL	80 dB SPL
Repetition rate	11.1/s &90.1/s	1.1/s
Sweeps	1500	250
Polarity	Rarefaction	Alternating
Electrode montage	Vertical	
Electrode sites	Inverting - Ipsi mastoid Non inverting - Cz (Vertex) Ground - Contra mastoid	
Amplification	1,00,000 times	50,000 times
Analysis time	12 ms	799 ms
Filters	100-3000 Hz	1-30 Hz
Notch filter	On	
Number of repetitions	2	

maintained within 5k Ω and relative impedances within 2 kΩ. Children who obtained normal wave V latency were only subjected to ACC recording. Protocol used for ABR and ACC recordings are shown in Table 1.

The consonant-vowel (CV) syllable /si/, spoken by an adult male native Kannada speaker was used as stimulus for ACC recording. The stimulus was recorded in a sound treated room using a dynamic microphone, placed at a distance of 10 cm from the lips of the speaker, at a sampling frequency of 44.1 kHz and 16 bit digitization. The recording and analysis of the stimulus was done using Adobe Audition software (version 1.5). Waveform of the syllable /si/, used in the study is shown in Figure 1. The duration of consonant portion /s/ was 149.6 ms, consonant vowel boundary was 2 ms, transition duration was 65.4 ms, Vowel duration (steady portion /i/) was 157 ms and total duration of /si/ was - 372 ms.

There were 2 conditions of stimuli presentation during ACC recording- /si/ syllable in quiet and /si/ syllable in noise (white noise presented ipsilaterally at 0 dB SNR). Totally, 4 recordings (2 in quiet & 2 in noise) were considered for analysis from each ear. A representative waveform of ACC recording with its latency and amplitude measures is shown in Figure 2.

Response Analysis

The replicable waves in each condition were averaged and analyzed for latencies and amplitudes by two experienced audiologists. Peaks were identified visually. Second positive peak of first LLR - P2, first negative peak, N1'and second positive peak, P2' of second LLR were marked. All latencies were calculated in milliseconds (ms). Also, peak to peak amplitude of P2-N1' and N1'-P2'complexes were calculated in microvolt (μV).

Results and Discussion

Overall data consisted of, Behavioral measure - SIN scores and Electrophysiological measures - Peak latency in ms (P2, N1' & P2'), and peak to peak amplitude in μV (P2-N1' & N1'-P2'). The data was tabulated and subjected to statistical analysis using the software, Statistical Package for the Social Sciences, SPSS (version 18).

Comparison of Measures between Ears

Paired t-tests were used to find differences between ears. Results indicated no significant differences ($p>0.05$). Hence, for further analysis measures of the two ears were combined.

Comparison of Stimulus Conditions (Quiet vs. Noise) and Groups (I vs. II)

Descriptive statistics of all latency and amplitude measures, for both groups in quiet and noise conditions are

Table 2: Mean and Standard Deviation of ACC Latencies and Amplitudes in Quiet and Noise

Parameters			Quiet		Noise	
			Group 1	Group 2	Group 1	Group 2
Latencies (ms)	P2	Mean	187.73	217.72	213.95	253.97
		SD	5.72	5.07	5.05	7.40
	N1’	Mean	262.38	283.09	286.96	322.74
		SD	4.80	4.92	6.55	7.27
	P2’	Mean	331.41	354.38	349.20	371.46
		SD	13.47	7.72	6.49	8.89
Amplitude (μV)	P2 -N1’	Mean	5.84	3.23	3.43	1.51
		SD	0.99	0.58	0.66	0.59
	N1’-P2’	Mean	2.46	1.58	2.02	1.42
		SD	0.67	0.43	0.58	0.36

shown in the Table 2. It can be observed from the tables that, when conditions were compared irrespective of the group, latency and amplitude measures were affected by noise i.e. latencies were delayed and peak to peak amplitudes were noticeably reduced. And when groups are compared, Group II (children with poor SIN scores) showed delay in latencies and reduction in amplitude compared to Group I (children with good SIN scores) in both the stimulus conditions.

To find whether these findings were statistically significant, Mixed ANOVA was performed. This provided information about main effects of conditions and groups and also the interactions between them. It was found that all main effects were significant in all the latencies and amplitudes ($p<0.001$). Also, interaction effects were significant except for the latency P2’ and peak to peak amplitude N1’-P2’. Reasons for such findings for conditions and groups are discussed in the following sections.

Comparison of Groups (I vs. II) within Each Latency and Amplitude

As the results of mixed ANOVA showed significant interactions between groups and conditions, Multivariate Analysis of Variance (MANOVA) was performed for comparing the groups within each latency and amplitude measure.

Latency measures: Figures 3 and 4 show means and standard deviations for latencies between groups in quiet and noise conditions respectively.

Results revealed that there is a significant difference between two groups in all the three latencies for both the stimulus conditions. Such results were obtained for brainstem responses in study by Anderson and Kraus (2010). They reported that noise induces latency delays in children with poor speech in noise perception. They attributed the reason for such findings to temporal processing deficits in the auditory brainstem of such children.

However, precise neuro-anatomical and neuro-physiological differences in the cortex of children with perceptual difficulties in noise are not explored still. Such differences might possibly disrupt the neural synchrony in those children which might be further degraded by adverse external conditions like noise leading to timing delays in cortical responses.

Amplitude measures: Figures 5 and 6 show means and standard deviations for amplitudes between groups in quiet and noise respectively.

Results showed a significant difference between two groups in both amplitudes for both the stimulus conditions. Anderson, Skoe, Chandrasekaran and Kraus (2010), studied brainstem correlates of speech in noise perception and found that children with poor perception in noise have reduced amplitude of neural measures. Explanation to such findings is majorly related to neuro-physiological differences of efferent system in good and poor listeners. Individuals with poor speech in noise scores have been speculated to have poor auditory efferent function (Kumar & Vanaja, 2004).

Studies on cortical potentials have failed to attribute the reason for decreased amplitude to any such central dysfunction observed in poor listeners. However, Anderson, Chandrasekaran, Yi and Kraus (2010), speculate that individuals with poor scores might be recruiting lesser neural resources due to lesser efficiency of the cortical pathway resulting in lesser amplitude.

Within Group Comparison of Latencies and Amplitudes between Conditions

Latency measure: Paired t test was used to study the latency differences between conditions within each group. Figures 7 and 8 show mean and standard de-

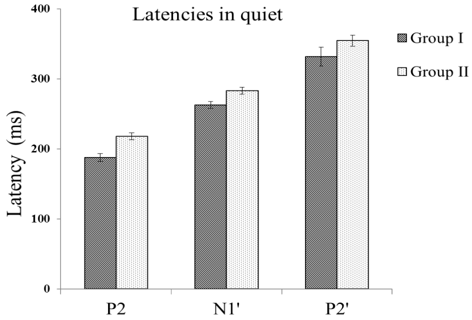


Figure 3: Means and Standard Deviations of latencies between Group I and II in Quiet.

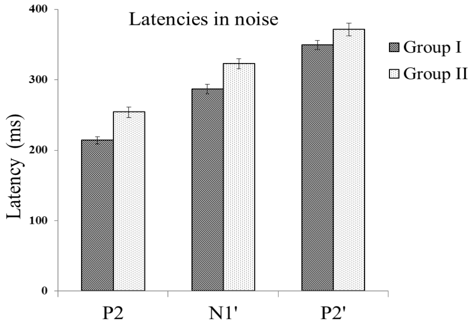


Figure 4: Means and Standard Deviations of latencies between Group I and II in Noise.

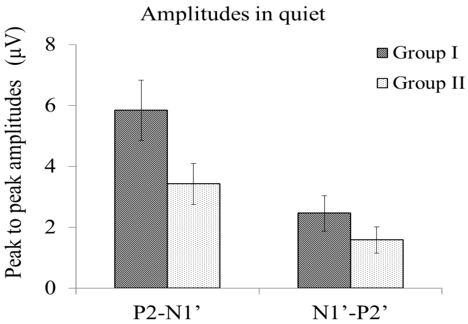


Figure 5: Means and Standard Deviations of amplitudes between Group I and II in Quiet.

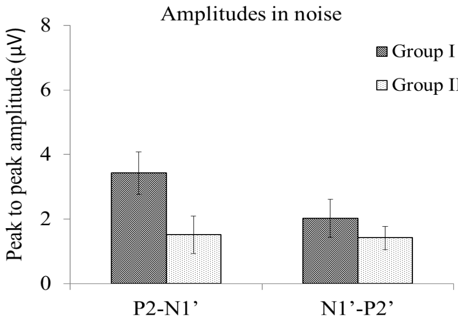


Figure 6: Means and Standard Deviations of amplitudes between Group I and II in Noise.

viations for latencies between stimulus conditions for groups I and II respectively.

Significant differences were obtained for all the comparisons of latencies between quiet and noise conditions in both the groups. These findings on latencies are in agreement with reports of several studies on evoked potentials both at sub-cortical (Cunningham et al., 2001) and cortical levels (Warrier, Johnson, Nicol & Kraus, 2004; Billings, Tremblay, Steker & Tolin, 2009). All those studies have reported delayed latencies for stimulus presented in noisy background. Possible reason for such findings can be disrupted neural synchrony due to noise which in turn may be reflected as delayed timing response of electrophysiological measures.

Amplitude measures: Paired t tests were used to study the amplitude differences between conditions within each group. Figures 9 and 10 show mean and standard deviation of amplitudes between stimulus conditions for group I and II respectively.

Again, significant differences were obtained for all the comparisons of amplitude measures between quiet and noise conditions in both the groups. As per the results, in both the groups introduction of noise has shown adverse effects resulting in reduced amplitudes. Wong, Uppunda, Parish and Dhar (2008) conducted a functional imaging study, where the stimuli were presented along with noise. Results revealed that cortical activation was less when stimuli were accompanied with noise and activation became lesser as the level of noise was increased. Such finding may support the speculation that amount of neural activation at the cortex is suppressed by noise which might result in reduced amplitudes of electrophysiological measures.

Other studies on cortical potentials reporting similar results are by Russo, Zeckler, Trommer, Chen and Kraus (2009) and Anderson et al. (2010). According to these researchers reduction in the amplitudes indicate poor sensory representation of acoustic aspects of signal in the presence of noise.

Correlation between Electrophysiological and Behavioral Measures

Pearson's correlation analysis was used to analyze the correlations between behavioral measure (SIN scores) and electrophysiological measures (latencies & amplitudes) in two conditions.

Results revealed a significant negative correlation ($p<0.001$) between latencies and SIN scores indicat-

Table 3: Correlation coefficient values between SIN Scores and Latencies in Quiet and Noise

Quiet			Noise		
P2	N1'	P2'	P2	N1'	P2'
-0.94	-0.91	-0.72	-0.94	-0.90	-0.76

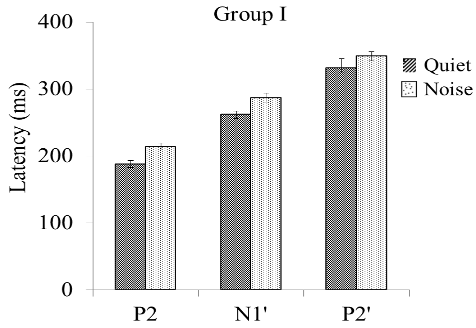


Figure 7: Means and Standard Deviations of Latencies between Conditions within Group I.

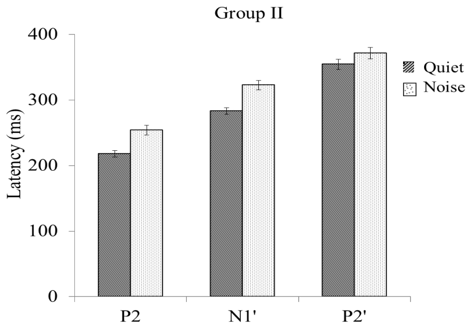


Figure 8: Means and Standard Deviations of Latencies between Conditions within Group II.

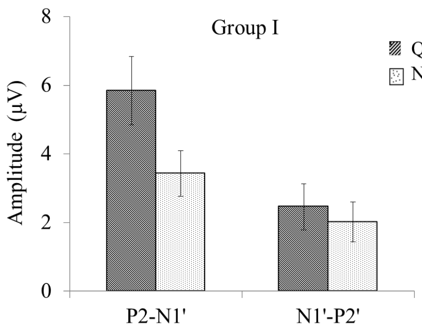


Figure 9: Means and Standard Deviations of amplitudes between conditions within Group I.

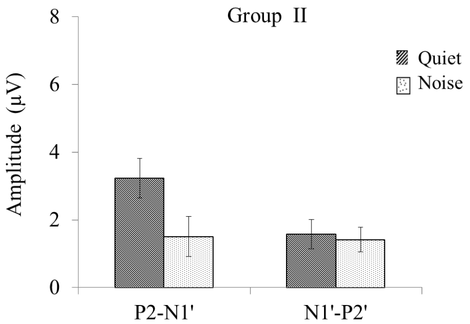


Figure 10: Means and Standard Deviations of amplitudes between conditions within Group II.

ing that as scores decrease the latencies increase and vice versa. Table 3 shows the correlation coefficient values between SIN scores and latencies. The above results show the robustness of relationship between cortical latency measures and speech in noise perception. These results are in support to the findings by Hornickel, Chandrasekaran, Zecker and Kraus (2011), who reported a link between behavioral SIN scores and subcortical neural measures. Results also showed a significant positive correlation ($p<0.001$) between amplitudes and SIN scores indicating that as scores decrease the amplitude also decrease and vice versa. Table 4 shows the correlation coefficient values between SIN scores and amplitudes. Anderson et al. (2010) correlated Hearing in Noise Test scores to N2 amplitude and found a significant correlation. Similarly, amplitude measures in this study also can be related to behavioral SIN scores. These results indicate that, electrophysiological measures act as neural signatures to behavioral speech perception in challenging environments.

Table 4: Correlation coefficient values between SIN Scores and Amplitudes in Quiet and Noise

Quiet		Noise	
P2-N1’	N1’-P2’	P2-N1’	N1’-P2’
0.84	0.61	0.82	0.61

Conclusions

The present study revealed that cortical potentials like ACC can also reflect the difficulties in speech in noise perception. Both latency and amplitude measures can be used to understand the perceptual and encoding difficulties posed by noisy background. These measures can potentially help us to identify children with difficulties in speech in noise perception. Robustness of the relationship between behavioral speech in noise and electrophysiological measures indicate that both at physical (as revealed by evoked potentials) and perceptual levels (as revealed by behavioral speech in noise test), effect of noise can be demonstrated. Also electrophysiological measures like latency and amplitude can act as neural correlates of speech in noise perception.

References

American National Standards Institute. (1999). *Maximum permissible ambient noise levels for audiometric test rooms* (ANSI S3.1-1999). New York: ANSI.

Anderson, S., & Kraus, N. (2010). Sensory-Cognitive Interaction in the Neural Encoding of Speech in Noise: A Review. *Journal of the American Academy of Audiology, 21,* 575-585.

Anderson, S., Chandrasekaran, B., Yi, H., Kraus, N. (2010). Cortical-Evoked Potentials Reflect Speech-in-Noise Perception in Children.*European Journal of neuroscience, 32,* 1407-1413.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *Journal of Neuroscience, 30*(14), 4922-4926.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27,* 32-38.

Billings, C. J., Tremblay, K. L., Stecker, G. C., & Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research, 254*, 15-24.

Boothroyd, A. (1997). Auditory development of the hearing child. *Scandinavian Audiology, 26,* 9-16.

Chermak, G. D., Vonhof, M. R., & Bendel, R. B. (1989).Word identification performance in the presence of competing speech and noise in learning disabled adults. *Ear & Hearing, 10*, 90-93.

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: Development and behavioral correlates of perception. *Ear and Hearing, 21,* 554-568.

Elliot, L. L., & Hammer, M. A. (1988). Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Disorders, 53,* 467-474.

Hnath- Chisolm, T. E., Laipply, E., & Boothroyd, A. (1998). Age related changes on a children's test of sensory-level speech perception capacity. *Journal of speech, language and hearing research, 41*, 94-106.

Hornickel J., Chandrasekaran, B., Zecker, S., & Kraus, N. (2011). Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. *Behavioural Brain Research, 216*, 597-605.

Jerger, J., & Jerger, J. (1970). Evoked responses to intensity and frequency change. *Archives of Otolaryngology, 91*, 433-436.

Kumar, A., & Vanaja, C. S. (2004). Functioning of olivocochlear bundle and speech perception in noise. *Ear and hearing, 25,* 142-146.

Martin, B. A. (2007). Can the acoustic change complex be recorded in an individual with a cochlear implant? Separating neural responses from cochlear implant artifact. *Journal of the American Academy of Audiology, 18*, 126-140.

Martin, B. A., & Boothroyd, A. (2000). Cortical, auditory, evoked potentials in response to changes of spectrum and amplitude. *Journal of the Acoustical Society of America, 107*, 2155-2161.

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

Ostroff, J. M., Martin, B. A., & Boothroyd, A. (1998). Cortical evoked responses to spectral change within a syllable. *Ear and Hearing, 19*, 290-297.

Ross, B., Tremblay, K. L., & Picton, T. W. (2007). Physiological detection of interaural phase differences. *Journal of the Acoustical Society of America, 121*, 1017-1027.

Russo, N. M., Zecker, S., Trommer, B., Chen, J., & Kraus, N. (2009).Effects of background noise on cortical encoding of speech in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 1185-1196.

Spoor, A., Timmer, F., & Odenthal, D. W. (1969). The auditory evoked responses to intensity modulated and frequency modulated tones and tone bursts. *International journal of Audiology, 8*, 410-415.

Treiman, R., Broderick, V., Tincoff, R., & Rodriguez, K. (1998). Children's phonological awareness: Confusions between phonemes that differ only in voicing. *Journal of Experimental Child Psychology, 68,* 3-21.

Vandana, S. (1998). *Speech identification test For Kannada speaking children.* Unpublished Independent project, University of Mysore, Mysore.

Warrier, C.M., Johnson, K.L., Hayes, E., Nicol, T., & Kraus, N. (2004). Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise.*Experimental Brain Research, 157*, 431-441.

Wong, P., Uppunda, A., Parrish, T., & Dhar, S. (2008). Neural basis of speech perception in noise. *Journal of Speech Language and Hearing Research, 51,* 1026-1041.

Comparison of Bone Anchored Hearing Aid with Test Band and Air Conduction Hearing Aid

[1]Vinsha K. & [2]Devi N.

Abstract

The aim of the present study was to compare the performance between the aided conditions with bilateral fitting of Bone Anchored Hearing Aid processors attached to test bands and binaural air conduction hearing aids, in bilateral conductive hearing loss. Fifteen individuals with bilateral symmetrical moderate to moderately severe degree of conductive loss were included in the study. Sound field warble tone thresholds, Speech Identification Scores in quiet as well as in the presence of noise and degrees of errors of horizontal localization were compared in the two aided conditions. Statistical analysis of data revealed that warble tone thresholds and Speech Identification Scores in quiet and noise were significantly better with bilateral Bone Anchored Hearing Aid processors attached to test bands compared to those with binaural air conduction hearing aids. There was no significant difference between the degrees of errors of horizontal localization in the two aided conditions.

Keywords:*Bone anchored hearing aid, Test band, Air conduction hearing aid.*

Introduction

Hearing loss can greatly affect the quality of life of an individual. It can have an impact on employment, education, and general well-being, unless and until it is properly managed. Fitting of air conduction hearing aids is considered to be an efficient treatment option for many individuals with hearing loss. However, it is contraindicated for patients with certain medical conditions such as recurrent otorrhoea, otitis media which is refractory to treatment, post operative anatomical deficits, congenital aural atresia and otitis externa (Bosman, Snik, Van der pouw, Mylanus & Cremers, 2001).

According to Spitzer, Ghossaini and Wazen (2002) the use of air-conduction hearing aids in persons with chronically draining ears entails risk of continuing or worsening infection caused by an earmold, which prevents adequate aeration of the ear. Eventhough venting is used in an effort to permit airflow and thus promote healing, often it results in feedback and inadequate gain. These venting efforts are often insufficient to allow substantial aeration, and thus the medical condition may be exacerbated. In addition, many persons with chronic otologic disease who have had prior ear surgery, such as a mastoidectomy, would have anatomical defects making air conduction hearing aid fitting a difficult task. The technical difficulties have been reported to include a challenging process of taking an impression in an ear with a mastoid bowl with risk of leaving material behind when the impression is removed. Having obtained an impression in such an ear, the fit may be problematic resulting in unmanageable feedback prohibiting significant hearing aid benefit. In these cases, one alternative option is the use of bone conduction devices for transmission of amplified sounds (Bosman et al., 2001), which bypass the normal sound passage through middle ear by vibrating the structures within the cochlea. A vibrator known as bone conductor is used as an output transducer. To effectively couple the vibrations to the skull and hence to the cochlea, the bone conductor is usually mounted on one side of a head band, which uses spring tension to push the bone conductor against the head. It can also be mounted on the arms of a spectacle aid. The hearing aid can be in a spectacle frame, in a BTE case mounted on the transducer headband, or in a body aid (Dillon, 2001).

Although conventional bone-conduction hearing aids have been used successfully for many years, they are associated with a number of practical problems resulting in limited use or patient rejection. Since an oscillator is held on the head using a headband and driven by a powerful hearing aid, it can result in discomfort caused by pressure on the mastoid which is crucial to deliver sufficient bone-conduction stimulation, but stretching of the band is common, leading to reduced sound quality and power. Frequent readjustments are usually required because of tension failures. Complaints of headache or ulcers involving the skin of the mastoid area may occur from the pressure against the skull (Spitzer et al., 2002). Maximum sound power output is limited due to acoustico-mechanical limitations of the transducer, limited static pressure, and the damping in the transmission path to the skull bone. Clinical practice shows that, due to the attenuation of the high frequencies by the skin and underlying tissue, sound quality is often judged rather poor when compared to air conduction aids. Finally, the static pressure necessary for correct operation of the aid by counteracting reactive forces often results in complaints of discomfort (Bosman et al., 2001).

In order to overcome the problems associated with both air- and bone-conduction hearing aids, the bone anchored hearing aids (BAHA) offers a reasonable alter-

[1]Email: vinsha.k@gmail.com,
[2]Lecturer in Audiology, Email: deviaiish@gmail.com

native. The BAHA takes advantage of the ability of bone to form a tight closure around a titanium implant. Attaching to a screw implanted into the mastoid, an abutment protrudes through the skin. The BAHA processor is snapped into place, eliminating the need for the headband and its side effects.

Bone anchoring utilizes a natural process called osseointegration. Osseointegration is the development of a solid connection between living bone and an implanted material (Chasin, 1999). Osseointegration of titanium implants was first demonstrated in the late 1960s by Per-Invar Branemark (Tjellstrom & Hakansson, 1995). The first clinical application of osseointegrated titanium implants was in the oral cavity to anchor a fixed bridge in an edentulous jaw (Branemark et al., 1977). Later in 1970s, Tjellstrom and his coworkers introduced the use of titanium implants outside the oral cavity for bone anchored hearing aids (cited in Spitzer et al., 2002).

Titanium is used for the implant screw, but research has also shown that some forms of stainless steel can also undergo osseointegration. In this procedure, a titanium screw is implanted into the temporal bone behind the ear. The osseointegration process takes approximately 3 months, after which the BAHA processor can be fitted on the patient. An external abutment is connected to the implanted screw, and the BAHA processor can be joined to this abutment with a simple bayonette connector (Chasin, 1999). The BAHA processor, consisting of microphone, amplifier and vibration transducer, can be connected and disconnected to the abutment by the wearer at will. Owing to the direct coupling to the temporal bone, BAHA has been proved to be superior in both wearer comfort and sound quality over conventional bone conduction hearing aids (Hakansson, Tjellstrom, Rosenhall & Carlsson, 1985).

Generally, for conductive or mixed hearing loss, the patient should have adequate sensorineural reserve measured by a bone-conduction curve of at least 45 dB HL for the head level processor, and an unaided speech discrimination score (word recognition score) greater than or equal to 60% (Habal, Frans, Zelski & Scheuerle, 2003). A bilateral fitting of BAHA should be considered for candidates with binaural hearing loss, which may lead to binaural hearing and in turn improving speech understanding, sound localization, and general candidate satisfaction (Hakansson, Tjellstrom & Rosenhall, 1984; Van der pouw Snik & Cremers, 1999; Grunder, Seidl, Ernst & Todt, 2008).

A preoperative assessment is recommended which includes sound field testing using a BAHA processor held in contact with the head by a special, bone-conduction-style headband or a soft, sweatband-style headband. Another measure is a test rod with a BAHA processor snapped into it. The test rod is held between the teeth with the mouth closed, allowing the patient to hear the conducted signal. In the use of test bands or test rod, there is some inefficiency of signal transduction, particularly in the high frequencies. Although none of these means of applying the BAHA mimic the post-implantation result precisely, this may assist in selecting the side to be implanted. Since it demonstrates the effectiveness of bone conducted stimulation, the experience is helpful to the patient in developing an understanding of the potential of BAHA (Spitzer et al., 2002).

Markides (1977), Festen and Plomp (1986), Day, Browning and Gatehouse (1988), Jerger, Darling and Florin (1995) have reported on the advantages of binaural application of air conduction hearing aids. Brooks (1984) assessed patient's subjective preference for monaural and binaural fitting and have shown that, in general binaural fitting was preferred. In contrast to air conduction hearing aids, only a few studies have been published on the advantage of binaural bone conduction device fitting. It has often been argued that the binaural application of any bone conduction device may not be effective due to the very less intracranial attenuation of skull vibrations leading to the stimulation of both the cochleae almost to the same extent (Beynon, Van der pouw, Mylanus & Cremers, 1998). However, Hamann, Manach and Roulleau (1991) reported that with bilateral application of bone anchored hearing aid, the speech reception threshold (SRT) in quiet was, on average, 4dB better than that with monaural application. However, any results on either sound localization or on speech recognition in noise was not included. Snik, Beynon, Van der pouw, Mylanus and Cremers (1998) studied sound localization and speech recognition in quiet as well as noise. The results revealed that there was an improvement in directional hearing for binaural bone anchored hearing aid application, but less directional hearing, or even none at all, for monaural application. Speech recognition threshold in quiet was found to be 3 to 6dB better with binaural bone anchored hearing aid and in the presence of noise there was an improvement of 2.9dB to 6dB with binaural fitting over monaural.

However, there is a dearth of studies on direct comparison between bilateral application of air conduction hearing aid and bone anchored hearing aids. Browning and Gatehouse (1994) suggested that pre-implantation evaluation of the difference in performance between the air-conduction hearing aid and a temporary conventional bone-conduction hearing aid might have value in predicting how patients who are advised to stop using their air conduction hearing aids will perform with BAHA. A point that must be noted here is that, on average patients perform significantly better with a BAHA than with a conventional bone-conduction hearing aid (Hakansson et al., 1990; Cooper, Burrell, Powell, Proops & Bickerton, 1996; Mylanus, Snik, Cremers, Jorritsma & Verschuure, 1994).

Thus the present study makes an attempt to compare the benefits of bilateral fitting of bone anchored hearing aid using test band and binaural fitting of air conduction hearing aid. The aim of the present study was to compare the performance with bilateral fitting of BAHA processors with test band and binaural air conduction hearing aids in individuals with bilateral conductive loss. The specific objectives were to compare the a) sound field warble tone thresholds with bilateral fitting of BAHA processors with test band and binaural air conduction hearing aids b) speech perception abilities in quiet as well as in the presence of background noise with bilateral fitting of BAHA processors with test band and binaural air conduction hearing aids and c) the horizontal localization abilities with bilateral fitting of BAHA processors with test band and binaural air conduction hearing aids.

Method

Participants

A total of 15 individuals with bilateral conductive hearing loss were included in the study. Age range of the participants was from 18 to 40 years. All participants had post-lingually acquired conductive hearing loss ranging from moderate to moderately severe degree with adequate speech and language. All the participants were oriented about the study and written consent was taken regarding their willingness to participate in the study. The participant selection criteria were as follows; Air-bone gap should be greater than or equal to 30 dB. Bone conduction thresholds should be less than or equal to 45dB. Air conduction thresholds and Bone conduction thresholds must be symmetrical (defined as less than 10 dB difference on average or less than 15 dB at individual frequencies) in both ears. Speech Recognition Threshold should be ?12 dB (re. PTA of 0.5, 1 and 2kHz). Word recognition should be greater than 60%. Age range was 18 to 40 years. Presence of middle ear pathology indicated by immittance evaluation. No indication of Retrocochlear Pathology (RCP). No history of neurological problems. No illness on the day of testing.

Testing Environment

All testing was carried out in a sound treated two room situation as per the standards of ANSI S3.1 (1999).

Instrumentation

A calibrated dual channel diagnostic audiometer, Madsen Orbiter 922 with TDH-39 headphones encased in MX 41AR ear cushion was used for performing the pure tone audiometry (air-conduction and bone-conduction) and speech audiometry in the unaided condition. The same audiometer with three Madsen loud speakers was used for performing speech identification tests in different aided conditions. One channel of the audiometer was connected to the loudspeaker placed at 0^0 azimuth. A toggle switch was used to route the signal of the other channel of the audiometer to any of the two speakers placed at $+45^0$azimuth or -45^0azimuth.

A calibrated GSI Tympstar (Version 2.0) middle ear analyzer was used to evaluate middle ear problems.

For evaluating the performance in aided conditions, four hearing aids were used; two digitally programmable air conduction hearing aids and two digitally programmable bone anchored hearing aids attached to head bands.

A personal computer with NOAH-3 and hearing aid specific software and the Hearing Instrument Programmer (HiPro) interface were used to program the digital Behind The Ear (BTE) air conduction hearing aids and digital Bone Anchored Hearing Aids (BAHA).

A laptop computer, installed with Adobe Audition software (version 3.0) was used to route the speech babble through the auxiliary input of the audiometer. Before the presentation of the stimuli, the level of the presentation was monitored with the calibration tone of 1 kHz. The level adjustment was manipulated in such way that it coincides with the 0dB in the audiometer's VU meter. The presentation level of the stimuli was monitored with the calibration tone. The same laptop was used to generate the stimulus for localization task. i.e, a train of white noise pulses, using Adobe Audition software (version 3.0).

For localization task, five Genelec 8020B loudspeakers mounted on Iso-Pod^{TM} (Isolation positioned/ Decoupler^{TM}) vibration insulating table stands were used. The loudspeakers were mounted at head level at five different angles. ie., at -90^0, -45^0, 0^0, $+45^0$ and $+90^0$keeping a distance of one meter from the patient's seat.

Cubase 6 software was used to present the localization stimulus from a personal computer. To route the stimulus to loudspeakers, Aurora 16 and Aurora 8 AD/DA converters were used. The output of the loudspeaker was calibrated using a sound level meter (Larson-Davis system 824, model no. 2540) with a 1/2" free-field microphone fitted to its preamplifier. The microphone of the sound level meter was placed at the position of the head of the participant, during calibration, at a distance of one meter. This process was carried out by presenting the stimuli through the loudspeakers, one at a time, and measuring the output for calibration. Thus, the loud speakers were calibrated to emit the output that would result in equal dB HL at the microphone at a distance of one metre.

Stimuli

Phonemically balanced (PB) word list in Kannada developed by Yathiraj and Vijayalakshmi (2005) was used for the measurement of Speech identification scores (SIS) in quiet and in the presence of noise. It consists of 4 lists, each having 25 words. Speech babble in Kannada developed by Anitha (2003) was used as background noise for the measurement of speech identification in noise. A train of four white noise pulses with duration of 200 ms separated by 200 ms of silence (Tyler et al., 2002) was generated for the purpose of localization task. A calibration tone of 1000 Hz was recorded prior to the train of white noise pulses. Stimulus was generated and normalized using Adobe Audition 3.0 software.

Procedures

The study was carried out in three phases; Selection of participants who have conductive hearing loss in both ears, Programming the air conduction hearing aids and BAHA and Comparison of sound field warble tone thresholds, Speech Identification Scores in quiet and noise and localization abilities.

Phase I. Selection of participants who have either conductive/ mixed hearing loss in both ears

Pure tone audiometric thresholds were estimated for air conduction at octave frequencies between 250 Hz and 8 kHz and bone conduction thresholds at octave frequencies between 250 Hz and 4 kHz using modified Hughson Westlake method (Carhart& Jerger, 1959). Speech audiometry was administered for all the participants in which Speech reception threshold, Speech identification scores and Uncomfortable loudness level for speech were found out.

Immittance evaluation using 226 Hz probe tone was carried out for all the participants. Tympanograms, ipsilateral and contralateral reflexes for stimulus frequencies of 500 Hz, 1 kHz, 2 kHz and 4 kHz were measured. Those individuals who met the participant selection criteria were included in the study.

Phase II. Programming the air conduction hearing aids and BAHA

Both air conduction hearing aids and digitally programmable BAHA processors were programmed using a personal computer and a HiPro interface unit using NOAH-3 and hearing aid specific fitting software.

The air conduction hearing aids were programmed to fit the hearing loss of the participant. NAL-NL1 fitting formula was used to prescribe the gain of the air conduction hearing aid according to the first fit.

BAHAs were programmed using specific fitting software for BAHA. The gain calculation was based on bone conduction thresholds. Additional gain at high frequencies was given as the present study assesses the pre-implantation evaluation of BAHA. It was intended to better approximate post-implantation results.

The hearing aid settings were optimized depending on participant's listening needs. Loudness normalization was done to make sure equal loudness in both ears in the aided conditions.

Phase III. Comparison of sound field thresholds, speech reception scores in quiet and noise and localization abilities

Testing was done in two aided conditions for each of the participants, namely aided condition with individually programmed air conduction hearing aids in both ears and aided condition with individually adjusted BAHA processors attached to test band on both the mastoids.

The following tests carried out in the above mentioned conditions were, Sound field thresholds for warble tones, Speech Identification Scores in four test conditions; quiet condition, Sound Front/Noise Front (SFNF) condition, Sound Front/Noise Right (SFNR) condition and Sound Front/Noise Left (SFNL) condition and Horizontal plane localization.

Sound field thresholds for warble tones: Sound field thresholds were obtained for warble tones at 500 kHz, 1 kHz, 2 kHz and 4 kHz. The warble tones were presented through loud speakers of the audiometer located at 0^o azimuth and at one meter distance from the participant. The minimum intensity at which the participant heard the warble tone 50% of the time were considered as the threshold. This procedure was carried out with the air conduction hearing aids in both the ears as well as with the BAHA processors attached to test band on both the mastoids which were individually programmed.

Speech Identification Scores in quiet: Speech Identification Scores in quiet were measured using PB word list in Kannada (Yathiraj & Vijayalakshmi, 2005). The participants were seated at a distance of one meter and at 0^o azimuth from the front loud speaker of the audiometer. The word list was presented using monitored live voice through microphone of the audiometer at 40dBHL. Speech Identification Score was measured for 25 words under each aided condition. The participants were instructed to repeat the words. A score of 1 was given for correct word repetition and a score of 0 was given for incorrect word repetition. The raw scores were converted to percentage scores by giving a weightage of 4% for each correct answer.

Speech Identification Scores in noise at 0dBSNR: To find out speech identification scores at 0dBSNR, the participants were seated at one meter distance at 0^o azimuth from the front loud speaker and one loudspeaker each was placed at 45^oazimuth on two sides. PB word list in Kannada (Yathiraj & Vijayalakshmi, 2005) was presented using monitored live voice at 40dBHL through front loudspeaker and speech babble was presented at the same level, through either the front, left or right loud speaker.There were three experimental conditions: Speech front/noise front (SFNF), Speech front/noise left (SFNL) and Speech front/noise right (SFNR).

Twenty five words were presented and the participants were instructed to repeat the words. A score of 1 was given for each correct word repetition and a score of 0 was given for each incorrect word repetition. The raw scores were converted to percentage scores by giving a weightage of 4% for each correct answer.

Horizontal plane localization: The participant was seated in the centre of the array of five loudspeakers. One loud speaker was placed in front of the patient at 0^{o} azimuth and two loudspeakers each to the right and left of the patient at $45^0$and 90^o azimuth.

A train of white noise pulses recorded on a compact disk was presented from a personal computer using Cubase 6 audio software and Aurora 16 and Aurora 8 AD/DA converters. Twenty five bursts of white noise were presented through the loudspeakers in a random order. The output of the loudspeaker was calibrated using a sound level meter with a free-field microphone fitted to its preamplifier.

A set of stimuli consisting of 25 similar trains of white noise pulses, five times from each loudspeaker, was presented in each of the two aided conditions (Bilateral BAHA with test band and binaural air conduction hearing aids). In each of the two aided conditions, 5 loudspeakers $\times$ 5 presentations, a total of 25, from each loud speaker were made. The stimuli were presented at 40 dBHL. During the test, the participants were instructed to maintain the designated position/orientation of the head. The order of 25 stimuli was randomized. The participants were instructed that he/she would be hearing a train of noise stimuli from any one of the five speakers at a time. Each time, he or she had to report the loudspeaker from which the stimulus was heard. The response mode from the participant was through a pointing task. The location of the loudspeaker to which participants pointed was noted down in terms of azimuth.

For the purpose of the study, Degree of error (DOE) was measured for the localization task. DOE corresponds to the difference in degrees between the degrees of azimuth of the loudspeaker of actual presentation of the stimuli, to the degree of azimuth of the loudspeaker identified as the source of the stimulus by the participant. For example, if the stimulus was presented from a loudspeaker at $+45^0$ azimuth and the participant reported the sound to be arriving from loudspeaker at -45^0, then the degree of error would be 90^0i.e., 45^0-(-45^0) = 90^0. This DOE was obtained for 25 trials in each aided condition. Thus, in each of the two different aided conditions, there was one set of degrees of errors consisting of 25 items.

A single representation of degree of errors in each aided condition was done by the calculation of root mean square degree of error (rms DOE) (Ching, Incerti, & Hill, 2004). The rms DOE is defined as the square root of the average of squared degrees of errors in each set. Thus, each participant had three rms DOEs, representing the localization abilities of the participants in the unaided condition and in each of the two aided conditions. It is calculated using the formula (Ching, Incerti, & Hill, 2004).

$$rmsDOE = \sqrt{\frac{DOE_1^2 + DOE_2^2 + DOE_3^2 + \cdots + DOE_{25}^2}{25}}$$

Where, DOE_n= Degree of Error of the n^{th} presentation in a set, and

$$rmsDOE = Root\ mean\ square\ degree\ of\ Error$$

The above data were tabulated and subjected to appropriate statistical analyses.

Results and Discussion

The results were tabulated and analyzed using the software SPSS version. 18.

Comparison of aided sound field threshold for warble tones in the two aided conditions

The mean and standard deviation (SD) of the sound field thresholds at 500, 1000, 2000 and 4000 Hz warble tones were obtained in the unaided and the two aided conditions. The mean and SD of these data are shown in the Table 1. To compare the warble tone thresholds obtained in the unaided condition and aided condition with bilateral BAHA processors attached to test bands, across frequencies, paired t-test was done. The result of paired t-test is given in Table 2.

The results revealed that the warble tone thresholds obtained in the unaided condition were significantly different from that obtained in the aided condition with bilateral BAHA processors attached to test bands.

Similarly, to compare the warble tone thresholds obtained in the unaided condition and the aided condition with binaural air conduction hearing aids, across frequencies, paired t-test was done and the result is given in Table 3.

To compare the warble tone thresholds obtained in the unaided condition and aided condition with bilateral

Table 1: Mean and Standard Deviation (SD) of the sound field thresholds for warble tones at different frequencies in the unaided condition and the two aided conditions

Condition	Warble tone detection thresholds across frequencies in dB			
	500Hz Mean (SD) dB HL	1kHz Mean (SD) dB HL	2kHz Mean (SD) dB HL	4kHz Mean (SD) dB HL
Unaided	51.00 (6.32)	49.00 (5.07)	46.00 (5.73)	43.33 (6.73)
Bilateral BAHA with test band	16.33 (3.99)	19.00 (5.73)	24.33 (6.23)	29.67 (3.99)
Binaural air conduction hearing aids	28.33 (7.94)	25.33 (8.12)	27.67 (5.94)	34.67 (7.90)

Table 2: Comparison of warble tone threshold across respective frequencies between the unaided condition and the aided condition with bilateral BAHA processors attached to test bands

Condition		Bilateral BAHA processors attached to test bands			
		500Hz	1kHz	2kHz	4kHz
Unaided	500Hz	**	-	-	-
	1kHz	-	**	-	-
	2kHz	-	-	**	-
	4kHz	-	-	-	**

*Note: - ** = Significantly Different at $p < 0.05$*

Table 3: Comparison of warble tone threshold across respective frequencies, between the unaided condition and the aided condition with binaural air conduction hearing aids

Condition		Binaural air conduction hearing aids			
		500Hz	1kHz	2kHz	4kHz
Unaided	500Hz	**	-	-	-
	1kHz	-	**	-	-
	2kHz	-	-	**	-
	4kHz	-	-	-	**

*Note: - ** = Significantly Different at $p < 0.05$*

BAHA processors attached to test bands, across frequencies, paired t-test was done. The result of paired t-test is given in Table 2.

The result of paired t-test revealed that the warble tone thresholds obtained in the unaided condition were significantly different from that obtained in the aided condition with binaural air conduction hearing aids.

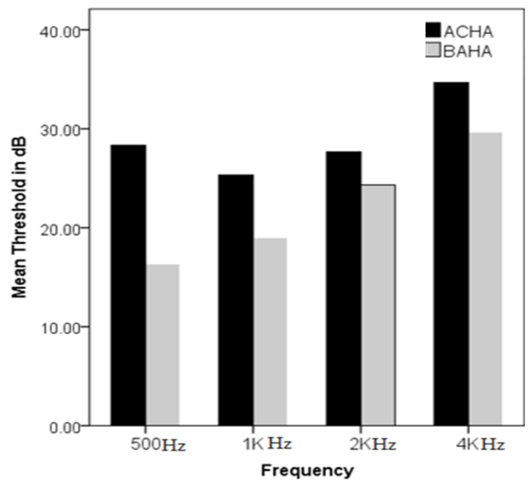


Figure 1: Mean warble tone thresholds obtained with bilateral BAHA processors attached to test bands and binaural air conduction hearing aids. Note: ACHA ? Binaural air conduction hearing aids BAHA ? Bilateral BAHA processors attached to test bands.

Figure 1 represents warble tone thresholds obtained with bilateral BAHA processors attached to test bands and binaural air conduction hearing aids. The mean warble tone thresholds with bilateral bone anchored hearing aid processors were lesser than that with binaural air conduction hearing aids. Paired t-test was done to find out whether these differences in mean threshold were statistically significant. The result of paired t-test, between the aided condition with bilateral BAHA processors attached to test bands and the aided condition with binaural air conduction hearing aids (the two aided conditions) is given in Table 4.

The results revealed that there was statistically significant difference in warble tone thresholds with the two aided conditions except at 2 kHz. In other words, the warble tone thresholds obtained with Bilateral BAHA processors were significantly better than those with binaural air conduction hearing aids at all frequencies except at 2 kHz.

Even though there was significant improvement with the aided condition with bilateral BAHA processors attached to test bands as well with binaural air conduction hearing aids compared to the unaided condition, the improvement with bilateral BAHA processors was significantly more in majority of the frequencies than with bilateral air conduction hearing aids. This can be due to the greater binaural loudness summation with bone conduction mode compared to that with air conduction mode. A possible reason for this is the differences in the interaural attenuation for these two modes of conduction, which varies from 0 to 15dB for bone conducted signals for octave frequencies from 250Hz to 4KHz. Whereas, the minimum interaural attenuation for air conduction signal is considered to be 40dB

Table 4: Comparison of warble tone threshold across respective frequencies between the two aided conditions

Aided condition		Bilateral BAHA processors attached to test band			
		500Hz	1kHz	2kHz	4kHz
Binaural air conduction hearing aids	500Hz	** -	-	-	-
	1kHz	-	**	-	-
	2kHz	-	-	**	-
	4kHz	-	-	-	**

*Note: - ** = Significantly Different at $p < 0.05$*

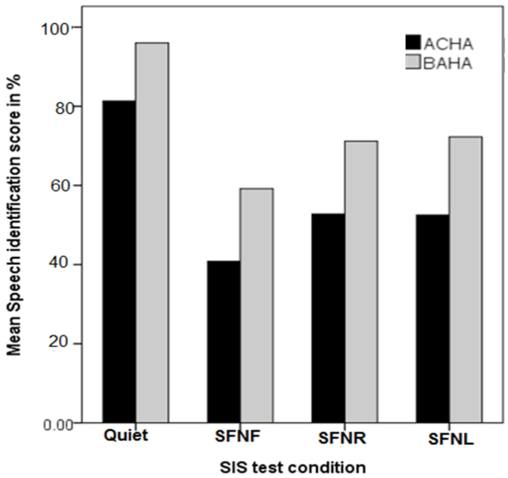


Figure 2: Mean warble tone thresholds obtained with bilateral BAHA processors attached to test bands and binaural air conduction hearing aids.

(Studebaker, 1967).

Another reason for the reduced threshold with BAHA processors at least in the low frequency can be the occlusion effect. Since the population considered for the present study is individuals with bilateral conductive hearing loss, the occlusion effect associated with the middle ear pathology, might have caused the louder perception of the bone conducted sounds (Roeser & Clark, 2007) through BAHA processors compared to the air conducted sound through air conduction hearing aids, leading to lower thresholds with binaural BAHA processors.

Comparison of Speech Identification Scores in Quiet and in the Presence of Noise in the Two Aided Conditions

Speech Identification Scores were obtained in four SIS test conditions. i.e, Quiet condition, Speech Front/Noise Front (SFNF) condition, Speech Front/Noise Right (SFNR) condition and Speech Front/Noise Left (SFNL condition. The mean and SD of Speech identification scores in the four SIS test conditions are given in Table 5.

To compare the Speech Identification Scores obtained in the unaided condition and aided condition with bilateral BAHA processors attached to test band, across the four SIS conditions, paired t-test was done. The result of paired t-test, between the unaided condition and the aided condition with bilateral BAHA processors attached to test band is given in Table 6.

The results showed that the Speech Identification Scores obtained in the unaided condition were significantly different from that obtained in the aided condition with bilateral BAHA processors attached to test band.

Similarly, to compare the Speech Identification Scores obtained in the unaided condition and the aided condition with binaural air conduction hearing aids, across the four SIS test conditions, paired t-test was done. The result of paired t-test, between unaided condition and the aided condition with binaural air conduction hearing aids is given in Table 7.

The results of paired t-test revealed that the Speech Identification Scores obtained in the unaided condition were significantly different from that obtained in the aided condition with binaural air conduction hearing aids.

To compare the speech identification scores obtained in the four different SIS test conditions using bilateral BAHA processors, one-way repeated measure ANOVA was done. The results revealed that there is significant difference in Speech Identification Scores across the four SIS test conditions at $p < 0.05$. Pair wise comparison was done using Bonferroni: Adjustment for multiple comparisons and the results of the test are given in Table 8.

The results showed that, with binaural BAHA processors attached to test band, there was no significant difference in speech identification scores between SFNR and SFNL conditions. That is, there was no signifi-

Table 5: The Mean and SD of Speech Identification Scores across the four SIS test conditions, in the unaided condition and the two aided conditions

Condition	Speech Identification Scores across the four SIS test conditions in %			
	Quiet Mean (SD)	SFNF Mean (SD)	SFNR Mean (SD)	SFNL Mean (SD)
Unaided	14.93 (13.81)	0.00 (0.00)	2.67 (4.70)	2.93 (4.65)
Bilateral BAHA with test band	95.60 (5.57)	59.20 (15.28)	70.93 (10.85)	72.27 (15.15)
Binaural air conduction hearing aids	81.33 (17.93)	40.80 (16.98)	52.27 (17.92)	52.00 (17.70)

Table 6: Comparison of warble tone threshold across respective frequencies between the two aided conditions

Aided condition		Bilateral BAHA processors attached to test band			
		Quiet	SFNF	SFNR	SFNL
Unaided	500Hz	** -	-	-	-
	1kHz		**	-	-
	2kHz	-	-	**	-
	4kHz	-	-	-	**

*Note: ** = Significantly Different at $p < 0.05$*

Table 7: Comparison of Speech Identification Scores across the four SIS test conditions, in the unaided condition and the aided condition with binaural air conduction hearing aids

Aided condition		Binaural air conduction hearing aids			
		Quiet	SFNF	SFNR	SFNL
Unaided	Quiet	**	-	-	-
	SFNF	-	**	-	-
	SFNR	-	-	**	-
	SFNL	-	-	-	**

*Note: ** = Significantly Different at $p < 0.05$*

Table 8: Pair wise Comparison across different SIS test conditions in the aided condition with bilateral BAHA processors

SIS test condition	Aided condition - Bilateral BAHA processors			
	Quiet	SFNF	SFNR	SFNL
Quiet	-	**	**	**
SFNF	**	-	**	**
SFNR	**	**	-	*
SFNL	**	**	*	-

*Note: ** = Significantly Different at $p < 0.05$*
*Note: * = Not Significantly Different at $p > 0.05$*

Table 9: Pair wise Comparison across different SIS test conditions with binaural air conduction hearing aids

SIS test condition	Aided condition - Bilateral BAHA processors			
	Quiet	SFNF	SFNR	SFNL
Quiet	-	**	**	**
SFNF	**	-	**	**
SFNR	**	**	-	*
SFNL	**	**	*	-

*Note: ** = Significantly Different at $p < 0.05$*
*Note: * = Not Significantly Different at $p > 0.05$*

cant difference between the speech identification scores when the noise came from left or right.

Speech identification scores were found to be significantly different between all other pairs of speech and noise conditions. From the mean data, it can be concluded that Speech identification scores obtained in

Table 10: Comparison of Speech Identification Scores obtained with the two aided Conditions across the four SIS test conditions

Aided condition		Binaural air conduction hearing aids			
		Quiet	SFNF	SFNR	SFNL
Bilateral BAHA processors	Quiet	**	-	-	-
	SFNF	-	**	-	-
	SFNR	-	-	**	-
	SFNL	-	-	-	**

*Note: ** = Significantly Different at $p < 0.05$*

quiet was better than that obtained in the presence of noise.

In the presence of noise, scores obtained in SFNR and SFNL were significantly better than that obtained in SFNF condition. In other words, better Speech Identification Scores were obtained when speech and noise came from different directions i.e, Speech from front and noise from either right or left direction, compared to the condition in which both speech and noise came from the same direction.

Similarly, to compare the speech identification scores obtained in the four different SIS test conditions using binaural air condition hearing aids, one-way repeated measure ANOVA was done. The results revealed that there is significant difference in Speech Identification Scores across the four SIS test conditions at $p < 0.05$. Pair wise comparison was done using Bonferroni: Adjustment for multiple comparisons and the results of the test are given in Table 9.

The results showed that, with binaural air condition hearing aids, there was no significant difference in speech identification scores between SFNR and SFNL conditions. That is, there was no significant difference between the speech identification scores when the noise came from left or right.

Speech identification scores were found to be significantly different between all other pairs of different SIS test conditions. From the mean data, it can be concluded that Speech identification scores obtained in quiet was better than that obtained in the presence of noise.

In the presence of noise, scores obtained in SFNR and SFNL were significantly better that that obtained at SFNF condition. In other words, better speech identification scores were obtained when speech and noise came from different directions i.e, Speech from front and noise from either right or left direction, compared to the condition in which both speech and noise came from the same direction.

Thus, across four different SIS test condition, both bilateral BAHA attached to test band and binaural air conduction hearing aids showed the same trend. That is, as expected, the Speech Identification Scores obtained in quiet condition were significantly better than that obtained with any other SIS test conditions.

In the presence of noise, scores obtained with SFNR and SFNL were significantly better than that obtained in SFNF condition. This is because, in the SFNF condition, since both the speech and noise came from the same direction, it would be very difficult to separate speech and noise. In SFNR and SFNL conditions, binaural unmasking might have played a role. It is due to binaural unmasking, a signal is detected in noise when interaural difference cues help the listener to isolate the signal from the noise (such as when the signal and the noise originate from different locations), as opposed to when there are no useful interaural difference cues (such as when only one ear is used or when the signal and noise originate from the same location). Since the speech came from front and noise came from right and left for the SFNR and SFNL conditions respectively (speech and noise came from different directions), the participants could make use of interaural cues to separate speech and noise. This finding is in accordance with the study done by Bronkhorst and Plomp (1988), in which they reported an improvement in intelligibility of speech as the interfering noise was moved away from the target speech location. They attributed to the fact of binaural unmasking and better ear listening.

Figure 2 represents the mean Speech Identification Scores in percentage, across four SIS test conditions. The mean Speech identification scores with bilateral bone anchored hearing aid processors were better than that with bilateral air conduction hearing aids. Paired t-test was done to find out whether these differences in mean were statistically significant. The result of paired t-test is given in Table 10

The results revealed that there was significant difference in Speech identification scores obtained with the two aided conditions across different SIS test conditions. From the mean data given in Table 5, it can be understood that Speech identification scores obtained with bilateral BAHA processors were significantly better in all conditions compared to binaural air conduction hearing aids.

The better speech perception in noise with bilateral BAHA processors can be due to the lesser distortion, because the BAHA processors as they bypasses the outer and middle ear and directly stimulate cochlea, very less gain is required. Whereas, additional gain had to be given for air conduction hearing aids so as to compensate for the conductive component or air-bone gap. As the amount of air-bone gap increases the amount of gain for air conduction hearing aids also has to be increased (Mylanus, van der Pouw, Snik & Cremers, 1998). Since all the participants considered for the present study had bilateral conductive hearing loss of more than 40dB, significantly more gain had to be increased for air conduction hearing aids compared to the very little gain needed for BAHA processors. The lesser distortion associated with the lesser gain and better loudness summation might have helped the participants to perform better with binaural BAHA processors.

Comparison of Horizontal Localization Skills in the Two Aided Conditions

The rms Degrees of error (DOE) of localization in the unaided condition and in the two aided conditions were found out and the mean and standard deviation (SD) was calculated. The mean and SD of this data are shown in the Table 11.

Paired t-test was done to compare the rms DOE in the unaided condition and that in the two aided conditions. The result showed that there was significant difference between the DOE of localization in the unaided condition and that with bilateral BAHA processors as well as with binaural air conduction hearing aids. Subjectively, participants reported that they felt more confusion in localization after wearing the aids, especially with 45^0 and 90^0azimuth. The mean data in Table 12 shows that, the mean rms DOE in the unaided condition was lesser than that in either of the aided conditions. This finding is similar to the findings by Van den et al. (2006). They reported that the localization ability of hearing-impaired listeners wearing hearing instruments has been shown to be worse than when not wearing hearing instruments.

Heyes and Ferris (1975) also reported that the localization performance by individuals with hearing loss was good with binaural postaural hearing aids. But it was still much inferior to the localization abilities of individuals with normal hearing.

The poorer performance in localization in both the aided conditions compared to the unaided conditions might be due to the disruption of Interaural Time Difference cues by small differences in signal processing on bilaterally worn devices, and distortion of Interaural Level Differ-

Table 11: The Mean and Standard Deviation (SD) of rms Degrees of Error (DEO) of localization obtained in the unaided and the two aided conditions

Condition	rms Degrees of Error Mean (SD)
Unaided	15.47 (18.89)
Bilateral BAHA with test band	32.45 (18.63)
Binaural air conduction hearing aids	32.65 (12.87)

ence by compression. Another possible explanation can be the microphone positions. For, both BAHA processors and air conduction hearing aids, the microphone position is behind the pinna resulting in obscured spectral information. This also might have led to localization confusions (Groth & Laureyns, 2011).

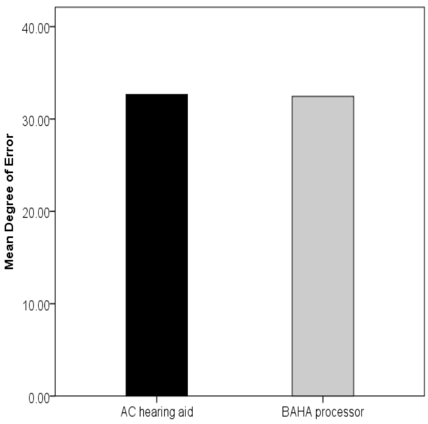


Figure 3: Mean rms Degrees of error of localization obtained with bilateral BAHA processors attached to test bands and binaural air conduction hearing aids.

The stimuli used for localization experiment were white noise bursts presented at 45dBHL which were audible to all of the participants even in the unaided condition. Since all of them had bilateral symmetrical hearing loss, significant localization difficulties were not present in the unaided condition.

Figure 3 represents the mean rms DOE of localization obtained in the two aided conditions. Paired t-test was done to compare the rms DOE values in the two aided conditions. The result showed that there was no significant difference in rms DOE obtained in the two aided conditions with $p > 0.05$. Thus, even though the localization skills with BAHA was under debate, because of the very less interaural attenuation of sounds leading to very limited interaural cues (Beynon et al., 1998), the results of the present study shows that the localization abilities with bilateral BAHA processors and that with binaural air conduction hearing aids are not significantly different.

Conclusions

From the present study it can be concluded that, the bilateral BAHA processors provide significantly better warble tone thresholds than binaural air conduction hearing aids. Also, the Speech identification Scores obtained with bilateral BAHA processors will be significantly better than that with binaural hearing aids, both in quiet and in the presence of noise. The Speech Identification Scores will be significantly better when speech and noise will be from different directions (SFNR and SFNL conditions) than when both were from the same directions in both the aided conditions. Further, no significant difference will be obtained in the rms degrees of errors of localization between the two aided conditions.

The study provides a support for bilateral implantation of BAHA in individuals with bilateral conductive hearing loss.

Also, it highlighted the better speech perception abilities with bilateral BAHA processors compared to bilateral air conduction hearing aids, both in quiet and in the presence of noise.Further,the results of the present study resolved the conflicts related to expected localization difficulties with bilateral BAHA due to the reduced intracranial attenuation.

Here are some future directions for research; Comparative study can be done with bilateral BAHA processors and binaural air conduction hearing aids in individuals with mixed hearing loss, the same study can be done grouping individuals with different amounts of air-bone gap and also localization experiments can be done with a low frequency and a high frequency stimulus as the effects of interaural time difference and interaural level differences can be studied.

References

American National Standards Institute (1999). *Maximum permissible ambient noise levels for audiometric rooms.* (ANSI S 3.1-1999). New York: ANSI.

Anitha, R. (2003). *The effect of speech babble of different languages on speech identification scores.* Unpublished independent project, University of Mysore, Mysore.

Beynon, A. J., Van der pouw, C. T. M., Mylanus, E. A. M., & Cremers, C. W. R. J. (1998). Binaural application of the bone anchored hearing aid. *Annals of otology, Rhinology & Laryngology, 107*, 187-193.

Bosman, A.J., Snik, A. F. M., Van der pouw, C. T. M., Mylanus, E. A. M., & Cremers, C. W. R. J. (2001). Audiometric Evaluation of Bilaterally Fitted Bone-anchored Hearing Aids. *Audiology, 40*, 158-167.

Branemark, P. I., Hansson, B. O., Adell, R., Breine, U., Lindstrom, J., Hallen, O., & Ohman A (1977). Osseointegrated implants in the treatment of the edentulous jaw. Experience from a 10-year period. *Scandinavian Journal of Plastic and Reconstructive Surgery, 16,* 1-132.

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America, 83*, 1508-1516.

Brooks, D. N. (1984). Binaural benefit. When and how much? *Scandinavian journal of Audiology, 13*, 237-241.

Browning, G. G., & Gatehouse, S. (1994). Estimation of the benefit of bone-anchored hearing aids. *Annals of Otology, Rhinology and Laryngology, 103*, 872-878.

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

Chasin, M. (1999). Update on implants: Bone-anchored devices and middle ear implants. *The Hearing Journal, 52*(7), 10-16.

Ching, T. Y. C., Incerti, P., & Hill, M. (2004). Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. *Ear and Hearing, 25*(1), 9-21.

Cooper, H. R., Burrell, S. P., Powell, R. H., Proops, D. W., & Bickerton, J. A. (1996). The Birmingham bone anchored hearing aid programme: referrals, selection, rehabilitation, philosophy and adult results. *Journal of Laryngology and Otology, 110*(suppl. 21), 13-20.

Day, G. A., Browning, G. G., & Gatehouse, S. (1988). *British journal of Audiology, 22,* 273-277.

Dillon, H. (2001). CROS, bone conduction and implanted hearing aids. *Hearing aids* (2nd ed.) (pp. 434-450). Australia: Boomerang press.

Festen, J. M., & Plomp, R. (1986). Speech reception threshold in noise with one or two hearing aids. *Journal of Acoustical Society of America, 79*, 465-471..

Groth, J., & Laureyns, M. (2011). Preserving localization in hearing instrument fittings. *The Hearing Journal, 64*, 34-38.

Grunder, I.,Seidl, R. O.Ernst, A., & Todt, I. (2008). Relative value of BAHA testing for the postoperative audiological outcome. *HNO, 56*, 1020-1024

Habal M., Frans, N., Zelski, R., & Scheuerle, J. (2003). Percutaneous bone-anchored hearing aid.*Journal of Craniofacial Surgery, 14*(5), 637-642.

Hakansson, B., Liden, B., & Tjellstrom, A, Ringdahl, A., Jacobsson, M., & Carlsson, P. et al. (1990). Ten years of experience with the Swedish bone-anchored hearing system. *Annals of Otology, Rhinology and Laryngology, 99*(suppl 151), 1-16.

Hakansson, B., Tjellstrom, A., & Rosenhall, U. (1984). Hearing thresholds with direct bone conduction versus conventional bone conduction. *Scandinavian Audiology, 13*(1), 3-13.

Hakansson, B., Tjellstrom, A., Rosenhall, U., & Carlsson, P. (1985). The bone anchored hearing aid: principle of design and a psychoacoustical evaluation. *Acta Otolaryngologica, 100,* 229-239.

Hamann C., Manach, Y., & Roulleau, P., (1991). Bone anchored hearing aid. Results of bilateral applications. *Revue de Laryngologie Otologie Rhinologie 112*(4), 297-300.

Heyes, A. D., & Ferris, A. J. (1975). Auditory localization using hearing aids. *British Journal of Audiology, 9*, 102-106.

Jerger, J., Darling, R., & Florin, E. (1995). Efficacy of cued listening task in the evaluation of binaural hearing aids. *Journal of the American Academy of Audiology, 5,* 279-285.

Markides, A. (1977). *Binaural hearing aids.* New York.; Academic press.

Mylanus, E. A. M., Snik, A. F. M., & Cremers, C. W. R. J. (1995). Patients' opinions of bone anchored vs conventional hearing aids. *Archives of Otolaryngology and Head and Neck Surgery, 121,* 421-425.

Mylanus. E. A. M., Snik, A. F. M., Cremers, C. W. R. J., Jorritsma, F. F., & Verschuure, J. (1994). Audiological results of the bone-anchored hearing aid HC200: multicenter results. *Annals of Otology Rhinology and Laryngolgy, 103*, 368-374.

Mylanus, E.A.M., van der Pouw, K.C.T.M., Snik, A. F.M., & Cremers, W.R.J. (1998). Intraindividual comparison of the bone-anchored hearing aid and air-conduction hearing aids. *Archives of Otolaryngology Head and Neck Surgery, 124*, 217-276.

Roeser, R. J., & Clark, J. L. (2007). Clinical masking. In Roeser, RJ., Valente, M., Dunn, H. H. (Eds.) *Audiology: Diagnosis* (2nd ed). 272-286, New York: Thieme.

Snik, A. F., Beynon, A. J., Van der pouw, C. T., Mylanus, E. A., & Cremers. (1998). Binaural application of the bone-anchored hearing aid. *Annals of Otology, Rhinology and Laryngology, 107*(3), 187-193.

Spitzer, J. B., Ghossaini, S. N., & Wazen, J. J. (2002). Evolving Applications in the Use of Bone-Anchored Hearing Aids. *American Journal of Audiology, 11*, 96-103.

Studebaker, G. A. (1967). Clinical masking of the non-test ear. *Journal of Speech and Hearing Disorders, 32*, 360-367.

Tjellstrom, A., & Hakansson, B. (1995). The bone anchored hearing aid: designs, principles, indications and long-term clinical results. *Otolaryngologic Clinics of North America, 28*(1), 53-72.

Tyler, R., Parkinson, A. J., Wilson, B. S., Witt, S., Preece, J. P., & Noble, W. (2002). Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. *Ear and*

Hearing, 23(2), 98-105.

Van den bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., & Wouters, J. (2006). Localization with bilateral hearing aids: without is better than with. *Journal of the Acoustical Society of America, 119,* 515-526.

Van der pouw, C. T. M., Snik, A. F. M., & Cremers, C. W. R. J. (1999). The BAHA HC200/300 in comparison with conventional bone conduction hearing aids. *Clinical Otolaryngology and Allied Sciences,24*(3), 171-176.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonemically balanced word list in Kannada.* Developed in Department of Audiology, AIISH, Mysore.

VEMP in Diabetes Mellitus

[1]Ghosh Vipin P.G. & [2]Sinha Sujeet Kumar

Abstract

Vestibulo ocular reflex and vestibule spinal reflexes are certain reflexes which are responsible for maintaining the balance and can be assessed by cervical Vestibular Evoked Myogenic potentials (cVEMP) and ocular VEMP (oVEMP) respectively. cVEMP reflects vestibular system activity that is elicited by high intensity sounds and detected as change in muscle potentials in the sternocleidomastoid muscle and oVEMP is a short latency vestibular evoked potential in response to a loud stimulus, recorded from contralateral extra ocular muscles. VEMP and oVEMP were recorded from thirty participants with and without diabetes mellitus. The latency and amplitude measures of cVEMP and oVEMP were studied within and across each group of participants. The findings revealed that the latency, amplitude and amplitude asymmetry ratio of peaks of cVEMP and oVEMP in participants without diabetes were in accordance to the previously published reports. Auditory Brainstem Responses in participants with diabetes mellitus were normal suggesting normal eighth nerve transmission till brainstem level. The duration of diabetes was not found to be a significant factor affecting pure tone average, cVEMP and oVEMP results. The latency of cVEMP and oVEMP peaks were not significantly different between individuals with and without diabetes mellitus but the amplitude of oVEMP and cVEMP responses were reduced in individuals with diabetes mellitus when suggesting a normal nerve conduction but a probable end organ deficit in persons with diabetes. The results also suggested that amplitude measures rather than latency measures were sensitive to vestibular end organ impairments.

Keywords: *cVEMP, oVEMP, Diabetes mellitus, duration of diabetes*

Introduction

Balance is dependent upon integration of signals from the vestibular, visual, and the somatosensory systems to generate the motor responses that maintain upright position and adjust to destabilizing forces. There are certain reflexes which are also responsible for maintaining the balance such as the vestibulo ocular reflex and vestibule spinal reflexes. The vestibular spinal reflexes and vestibule ocular reflexes can be assessed by cervical Vestibular Evoked Myogenic potentials (cVEMP) and ocular VEMP (oVEMP) respectively with similar threshold of elicitation. cVEMP reflects vestibular system activity that is elicited by high intensity sounds and detected as change in muscle potentials within the neck (Colebatch, Halmagyi & Skuse, 1994). oVEMP was first demonstrated by Todd, Rosengren and Colebatch, (2003).They demonstrated a short latency vestibular evoked potential with a negative peak at 10 ms (n10) and a positive peak around 15 ms (p15) in response to a loud 500 Hz bone-conducted stimulus, which could be best recorded from contralateral extra ocular muscles (Rosengren, Todd & Colebatch, 2005). These tests serve as an important tool in the test battery of clinics and medical centers worldwide offering services for patients with balance and vestibular disorders.

Among the many causes of vestibular symptoms diabetes mellitus is one. This systemic disorder alters the normal glucose metabolism leading to glycemia increase beyond physiological levels that causes glucose build up within bodily fluids (hyperglycemia) and changes the osmotic potential and the functioning of all the systems, the vestibular among them (Adriano & Silvia, 2006). Diabetes mellitus is a genetically determined metabolic disorder associated with absolute or relative impairment of insulin and in complete clinical manifestation is characterized by metabolic affections, vascular and neuropathic complications (Maia & Campos, 2005). Individuals with diabetes mellitus may also suffer from central neuropathy, or degeneration of the higher nervous system (Jadzinsky, Faerman & Fox, 1973; Shagan, 1976; Tavormina, Kastner, Slater & Watts, 1976).

Although the prevalence of vestibular deterioration is not well known in persons with diabetes mellitus, recent studies suggest that about 60% to 65% of patients with diabetes mellitus may have vestibular dysfunction, associated with deficits in gaze holding and the vestibular reflex (Jauregui- Renoud, Sanchez, Olmos & Golzalez-Barcena, 2009). In individuals with diabetes, problemsÂ in both the timing and quality of gait have been reported to be absent (Petrofsky,Cuneo, Lee,Johnson & Lohman, 2006; Petrofsky, Lee, Macnider & Navarro, 2005). However, there is dearth of information regarding the other various test findings in individuals with diabetes. Since most of the individuals with diabetes are asymptomatic, the detailed vestibular evaluation has not been done in individuals with diabetes.

This research aimed to study the latency and amplitude measures of cVEMP and oVEMP responses of individuals with diabetes, individuals without diabetes mellitus and compare between these two groups. Attempts were

[1]Email: vipinghosh78@gmail.com,
[2]Lecturer in Audiology, Email: sujitks5@gmail.com

made to correlate the findings of cVEMP and oVEMP, VEMP results with hearing threshold, VEMP results with duration of diabetes and VEMP results with the symptomatology presented by the participants.

Method

Participants

Two groups of participants with normal hearing sensitivity or sensorineural hearing loss up to moderate degree were considered for the study. History of conductive pathology or neuromuscular problems especially of that of head and neck region and a history of intake of drugs that lead to vestibulo toxicity were ruled out in all the participants.

Group 1-Experimental group: 30 participants (60 ears) in the age range of 35 to 55years with diabetes (mean age= 44.1 years) were selected if they had medically confirmed history of type-II diabetes mellitus and if there was no evidence of any retro cochlear pathology confirmed with Auditory Brainstem Response (ABR) results. These participants were screened for uncomfortable loudness level of $>$ 105dB for speech.

Group 2- Control group: 30 participants (60 ears) in the age range of 35 to 55 years without diabetes (mean age= 43.13 years) were selected if they had no history or presence of medically confirmed diabetes mellitus.

A calibrated two channel GSI-61 diagnostic audiometer connected to a TDH-39 headphone and Radio ear 71 bone vibrator was used to find out the air conduction thresholds, bone conduction thresholds and uncomfortable level for speech from all the participants. Calibrated GSI TYMPSTAR immittance meter was used for tympanometry and reflexometry. Intelligent Hearing systems (IHS version 4.3.02) was used for recording auditory brainstem responses and air conducted tone burst evoked cervical VEMP. Calibrated Eartone 3-A insert earphone was used to deliver the stimuli. Biologic navigator ProEP instrument with biologic insert was used for ocular VEMP recordings. The ambient noise levels inside the test room were within permissible limits (ANSI, 1999).

Test Procedure

A detailed case history was taken for each participant prior to testing. It was followed by administration of the five sections of dizziness questionnaire (Maryland Hearing and Balance Centre, 2004) for all the participants. Pure-tone thresholds were obtained for all the participants using modified Hughson and Westlake procedure (Carhart & Jerger, 1959) at octave frequencies between 250 Hz and 8000 Hz for air conduction and between 250 Hz and 4000 Hz for bone conduction. UCL was obtained in both ears for air conducted speech stimuli using ascending method. Immittance evaluation was carried out in both ears using a probe tone frequency of 226 Hz. Tympanometry was done initially and then ipsilateral and contralateral acoustic reflex thresholds were measured for 500, 1000, 2000, and 4000 Hz stimuli. For all the participants ABR were recorded for both the ears to rule out any retrocochlear pathology. ABRs were recorded using an electrode montage of non inverting electrode to the upper forehead, inverting to the ipsilateral (stimulated) ear lobe and ground to lower forehead. The amplifier band pass was 100 Hz to 3000 Hz. Alternating polarity click stimuli were presented monaurally at a rate of 11.1 Hz and 90.1 Hz at 90 dB nHL. Averaged responses to 1500 clicks were collected on each of two runs. Reproducible components were obtained and latencies were measured.

For the cVEMP recordings the sites of electrode placement were prepared using a skin preparation gel. Silver chloride disc electrodes were used for recording. Absolute electrode impedances and inter electrode impedances were maintained below 5000 ohms and 2000 ohms respectively. During the cVEMP recordings the participants were instructed to sit straight and turn their head to the opposite side of the ear in which stimulus was presented, so as to activate ipsilateral sternocleidomastoid muscle, as it gives reliable and greater amplitude. Participants were instructed to maintain the same posture throughout the test run. A visual feedback box in the IHS instrument with green and red LED lights was provided to the participant in order to maintain the tonicity of the sternocleidomastoid muscle within 50microVolt and 100 microvolt. cVEMPs are recorded using 500 Hz tone burst (2 cycles rise, 0 cycles plateau, and 2 cycles fall, Blackman weighting function) in rarefaction polarity which was presented at a rate of 5.1/s. ER - 3A insert ear phone was used to present the stimulus at an intensity of 95 dBnHL to each ear. A pre stimulus period of 10ms and post stimulus duration of 70ms was used to record the stimulus. The amplified responses (X 5000) were band pass filtered between 30 to 1500 Hz. The responses were averaged totally for 200 stimuli. Responses were recorded twice to ensure the replicability of the responses.

Ocular VEMP was recorded for all the participants with upper gaze direction. Participants were instructed to maintain the same upper gaze throughout the test run. The recordings were done from the extraocular muscles contralateral to the ear being stimulated. oVEMPs were recorded for all the participants with upper gaze direction. Participants were instructed to maintain the same upper gaze throughout the test run. Stimuli used to record oVEMPs were identical to stimuli used to record cVEMPs. 500 Hz tone burst (2 cycles rise, 0 cycles plateau, and 2 cycles fall, Blackman weighting function) was presented in rarefaction polarity at a rate of 5.1/s. ER - 3A insert ear phones was used to present the stimulus at an intensity of 95 dBnHL to each ear. A pre stimulus period of 10ms and post stimulus duration of 60ms was used to record the stimulus. The amplified responses (X 5000) were band pass filtered between 1 to 1000 Hz. The responses were averaged totally for 200 stimuli. oVEMP responses were recorded twice to ensure the replicability of the responses.

Response Analysis

The recorded cVEMP and oVEMP responses were analyzed for latency and amplitude measures of the peaks (cVEMP: P13, N23; oVEMP: n1, p1 and n2) for all the participants in the control and experimental groups. The amplitude asymmetry ratio for P13-N23 complex in cVEMP , n1-p1 and p1-n2 complex in oVEMP were calculated using the following formulae.

$$\text{Amplitude assymetry ratio} = \frac{\text{amplitude of P13-N23 complex in the right ear-amplitude of P13-N23 complex in the left ear}}{\text{amplitude of P13-N23 complex in the right ear+amplitude of P13-N23 complex in thge right ear}} \text{ X 100}$$

$$\text{Amplitude assymetry ratio} = \frac{\text{amplitude of n1-p1 complex in the right ear-amplitude of n1-p1complex in the left ear}}{\text{amplitude of n1-p1 complex in the right ear+amplitude of n1-p1complex in thge right ear}} \text{ X 100}$$

$$\text{Amplitude assymetry ratio} = \frac{\text{amplitude of p1-n2 complex in the right ear-amplitude of p1-n2 complex in the left ear}}{\text{amplitude of p1-n2 complex in the right ear+amplitude of p1-n2 complex in thge right ear}} \text{ X 100}$$

Note: For the ease of reading the P13 and N23 peaks have been written as P1 and N1 respectively throughout results and discussion.

Results

Results of Control Group

Latency and amplitude of cVEMP and oVEMP in control group: The peaks of cVEMP (P1 & N1) and oVEMP (n1, p1 & n2) were identified and latency, amplitude and amplitude asymmetry ratio was calculated from all the participants in the control group Figure 1 shows a cVEMP and oVEMP recorded from a participant in the control group.

The mean and standard deviation (SD) of latency measures of P1 and N1 peaks, amplitude of P1-N1 complex of cVEMP for right and left ears, similar results of oVEMP and amplitude asymmetry ratio for P1-N1 complex and n1-p1 and p1-n2 complex for participants in the control group is given in Table 1.

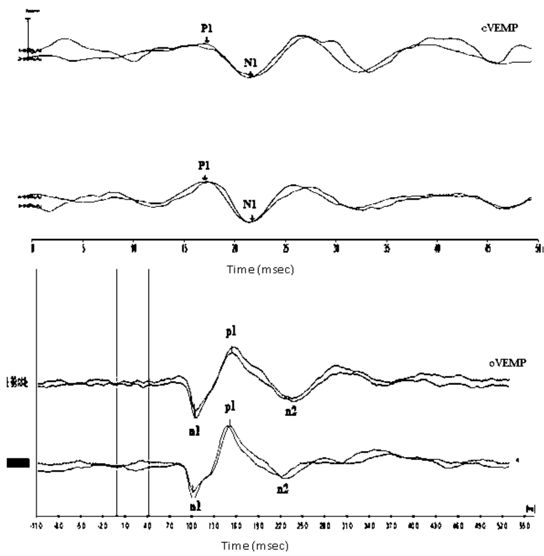


Figure 1: Patterns of cVEMP and oVEMP recorded from a participant in the control group.

From Table 1, the latency of peaks of oVEMP was early compared to the peaks of cVEMP. Also, the amplitude of cVEMP wa higher compared to that of oVEMP. The standard deviation for amplitude of both cVEMP and oVEMP was high in the control group.

Correlation of latency and amplitude measures of cVEMP and oVEMP in control group: Correlation of latency and amplitude measures across cVEMP and oVEMP was studied using Pearson's test of correlation. The cVEMP and oVEMP data from both ears were combined for statistical analysis. There was no significant correlation between the amplitude of P1N1 complex of cVEMP with amplitude of n1p1 complex of oVEMP ($r= -0.045$, $p > 0.05$) and p1n2 complex of oVEMP ($r=0.084$, $p > 0.05$), P1 latency of cVEMP with n1 latency of oVEMP ($r=0.211$, $p > 0.05$), P1 latency of cVEMP with p1 latency of oVEMP ($r=0.195$, $p > 0.05$) and P1 latency of cVEMP with n2 latency of oVEMP ($r=0.120$, $p > 0.05$). To conclude, there was no correlation between the latencies and amplitude of cVEMP

Table 1: Mean and standard deviation of latency measures, amplitude measures and amplitude asymmetry ratio of P1 & N1 peaks of cVEMP and n1,p1 and n2 peaks of oVEMP

	cVEMP							oVEMP												
	Latency measure (ms)				Amplitude measure (μV)		P1-N1 AmAR (%)	Latency measure (ms)						Amplitude measure (μV)				AmAR(%)	
	P1		N1		P1-N1			n1		p1		n2		n1p1		p1n2		n1p1	p1n2
	Right	Left	Right	Left	Right	Left		Right	Left	Right	Left	Right	Left	Right	Left	Right	Left		
Mean	14.80	14.75	21.37	21.89	38.3	38.64	18.64	11.97	12.16	17.10	17.38	23.13	22.73	5.59	6.21	4.58	5.28	23.28	27.86
SD	1.83	2.47	2.07	2.26	14.34	15.97	12.64	1.87	1.29	1.68	1.57	1.82	1.96	2.99	4.07	2.82	3.38	19.86	18.68

**Note: AmAR= Amplitude Asymmetry Ratio* with the latencies and amplitude of oVEMP for the control group.

Results of Experimental Group

Latency and amplitude of cVEMP and oVEMP: The mean pure tone threshold average (average of thresholds at 500Hz, 1000Hz and 2000Hz) was 19.26 (SD= 4.80) in the right ear and 19.44 (SD= 4.41) in the left ear. From the results of pure tone audiometry done on 60 ears, 45 ears (75%) had a pure tone threshold average above 15dBHL. Further, out of these 45 ears, 7 ears (11.6%) had pure tone threshold average above 25 dBHL (Mild hearing loss) lesser than 30dB. Normal immitance and acoustic reflex was obtained from all the participants in the experimental group. The ABR could be recorded from all the 60 ears and was found to be normal in its latency and amplitude measures.

The VEMP responses were not present in all the participants in experimental group. Out of the 60 ears tested in experimental group (30 participants with diabetes mellitus) for VEMP, cVEMP was absent in 48 ears (80%) and oVEMP was absent in 41 ears (68.33%). This reduced the data corpus for statistical analysis to 12 ears for the cVEMP and 19 ears for oVEMP in the experimental group. Further, in the experimental group, both the cVEMP and oVEMP were absent in 34 ears (56.6%), cVEMP was present and oVEMP was absent in 7 (11.6%) ears, cVEMP was absent but oVEMP was present in 14 (23.33%) ears. Both cVEMP and oVEMP were present only in 5 (8.3%) ears. The latency of P1 and N1 peaks of cVEMP and amplitude complex of P1-N1 peaks were measured for 12 ears. Figure 2 (a) and (b) shows cVEMP recordings from persons with diabetes- one with cVEMP response and another with no cVEMP response respectively. The latency of n1, p1 and n2 peaks of oVEMP was also obtained for 19 ears. Figure 3 (a) and 3 (b) shows the absent and abnormal recordings of oVEMP from persons with diabete. Descriptive analysis was done to find out the mean and standard deviation (SD) of latency and ampltiue measures of peaks of cVEMP in 12 ears and oVEMP in 19 ears of participants with diabetes (Table 2).

Correlation of latency and amplitude measures of cVEMP and oVEMP: Spearmans test of correlation revealed no significant correlation between latency of P1

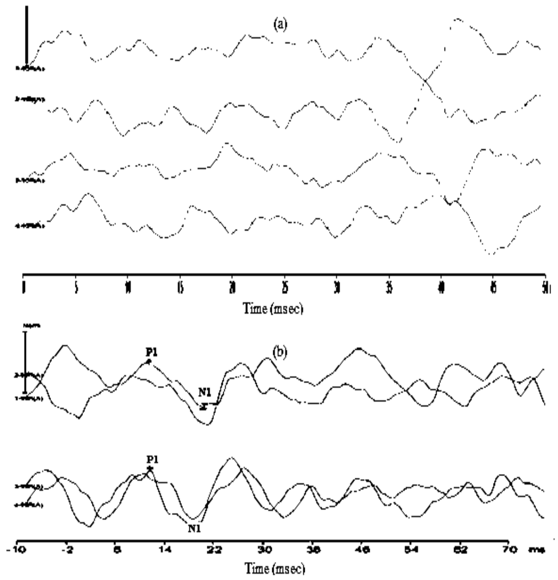


Figure 2: Patterns of cVEMP recorded from individual with diabetes (a) absence of cVEMP and (b) presence of cVEMP.

peak of cVEMP and n1 peak of oVEMP (p=0.370, p>0.05), P1 peak of cVEMP and p1 peak of oVEMP (ρ =0.074, p>0.05), P1 peak of cVEMP and n2 peak of oVEMP (ρ =0.361, p>0.05), N1 peak of cVEMP and n1 peak of oVEMP (ρ =0.353, p>0.05), N1 peak of cVEMP and p1 peak of oVEMP (ρ =0.032, p>0.05), N1 peak of cVEMP and n2 peak of oVEMP (ρ=0.032, $p > 0.05$) in the participants of experimental group. Similar results were obtained for amplitude measures of cVEMP and oVEMP (P1N1 of cVEMP and n1p1of oVEMP: ρ =-0.126, p>0.05, P1N1 complex of cVEMP and p1n2 complex of oVEMP: ρ =0.357, p >0.05).

Correlation of pure tone average and the results of cVEMP and oVEMP: The findings obtained for the cVEMP and oVEMP for participants in the experimental group were classified into three categories: normal, abnormal and absent. A Pearson test of correlation was done to study the correlation of pure tone average (average of hearing thresholds at 500 Hz, 1000 Hz and 2000 Hz tones) and the results of cVEMP and oVEMP in the participants of the experimental group. There was no significant correlation between pure tone average of right ear with cVEMP results in the right ear (r=-0.61, p >0.05), pure tone average of right ear with oVEMP

Table 2: Mean and standard deviation of latency and amplitude measures of cVEMP and oVEMP in participants of experimental group

	cVEMP			oVEMP				
	Latency measures (ms)		Amplitude measures (μV)	Latency measures (ms)			Amplitude measures (μV)	
	P1	N1	P1-N1	n1	p1	n2	n1-p1	p1-n2
Mean	14.91	20.93	26.31	11.85	17.50	22.95	3.82	4.08
SD	1.74	2.053	5.47	1.36	1.90	2.53	3.93	4.72

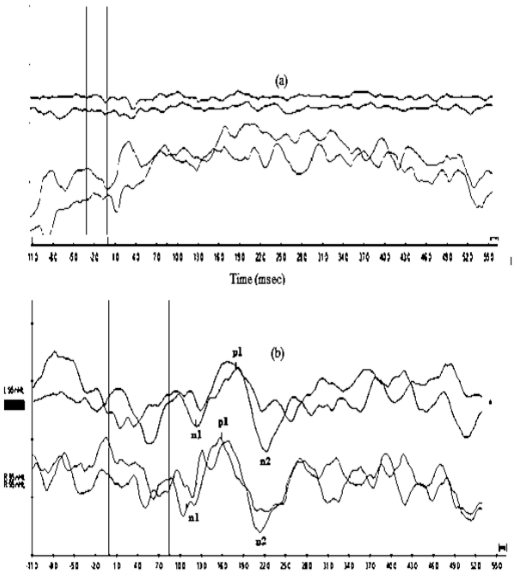


Figure 3: Patterns of oVEMP recorded from participants in the experimental group. (a). absence of oVEMP, and (b) Presence of oVEMP.

results in the right ear (r=-0.16, p >0.05), pure tone average of left ear with cVEMP results in the left ear (r=-0.13, p >0.05), pure tone average of left ear with oVEMP results in the left ear (r=-0.12, p >0.05).

Correlation of duration of diabetes mellitus and latency and amplitude measures of cVEMP and oVEMP: The mean duration of diabetes mellitus in the participants of experimental group was calculated and was found to be 12.53 ± 8.10 months. Spearmans test of correlation was run to study the correlation of duration of diabetes mellitus with latency and amplitude measures of cVEMP and oVEMP in experimental group of participants. No significant correlation was found between the duration of diabetes mellitus and latency of cVEMP (P1: ρ=0.35, p>0.05; N1: ρ =0.62 , p>0.05) and oVEMP (n1: ρ =-0.16, p>0.05; p1: ρ =0.08, p>0.05; n2: ρ =0.05, , p>0.05). There was no significant correlation of amplitude of cVEMP (P1-N1: ρ =0, p>0.05) and oVEMP (n1p1: ρ =-0.13, p>0.05, p1n2: ρ =-0.10, p>0.05) with the duration of diabetes.

Comparison between Control and Experimental Group

Results of cVEMP

The mean latency of P1 and N1 peak for control and experimental group is shown in Figure 4. From Figure 4, the latencies of P1 for both the control and the experimental group were similar whereas, the latency of N1 peak for experimental group was lesser compared to the control group. To statistically understand these measures, a non parametric Mann Whitney U test was carried out and it revealed no significant difference in the latency of P1 peaks (Z= -.378, $p > 0.05$) and N1 peaks (Z= -1.649, $p > 0.05$) across control and experimental groups.

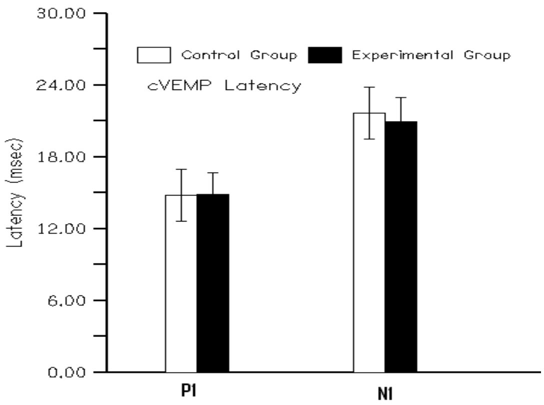


Figure 4: Mean and standard deviation of latency (in ms) of P1 and N1 peaks of cervical VEMP in control and experimental group.

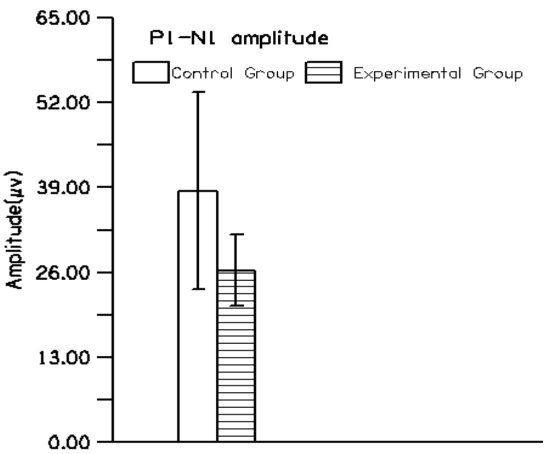


Figure 5: Mean and standard deviation of amplitude (in µV) of P1-N1 complex of cervical VEMP in control and experimental group.

Figure 5 depicts the mean amplitude measures of P1N1 complex. The amplitude of P1N1 complex for the experimental group was lesser compared to the control group. Mann Whitney U test revealed significant difference in the mean amplitude of P1N1 complex across groups (Z= -2.856, $p < 0.05$).

Results of oVEMP

The latency of n1, p1 and n2 peaks of oVEMP was calculated for both the control and the experimental groups (Figure 6). From the figure, the difference between control and experimental group in terms of latencies of oVEMP was not evident. Statistically significant differences in latency measures across the groups were studied with non parametric Mann Whitney U test and

Table 3: Percentage of individuals reported with various vestibular symptoms and their cVEMP results

	cVEMP right				cVEMP left			
	A	N	Ab	Total %	A	N	Ab	Total %
Light headedness or swimming sensation in the ear	0	2(50%)	2(50%)	4(13.3%)	3(75%)	1(25%)	0	4(13.3%)
Blacking out or loss of consciousness	0	1(100%)	0	1(3.3%)	0	1(100%)	0	1(3.3%)
Tendency to fall	10(83.3%)	0	2(16.6%)	12(40%)	13(92.8%)	0	1(7%)	14(46.6%)
Objects spinning or turning around you	7(100%)	0	0	7(23.3%)	8 (100%)	0	0	8(26.6%)
Sensation that you are turning or spinning inside	0	0	1(100%)	1(3.3%)	0	1(100%)	0	1(3.3%)
Loss of balance when walking	5(100%)	0	0	5(16.6%)	2(100%)	0	0	2(6.6%)
Total	22(73.3%)	3(10%)	5(16.6%)	30(100%)	26(86.6%)	3(10%)	1(3.3%)	30(100%)

Note: cVEMP= Cervical Vestibular Evoked Myogenic Potential; A= Absent, N= Normal, Ab= Abnormal

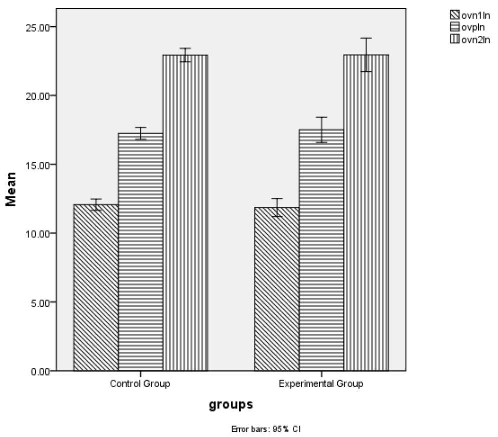


Figure 6: Mean and standard deviation of latency (in ms) of n1, p1 and n2 peaks of oVEMP in participants of control and experimental groups.

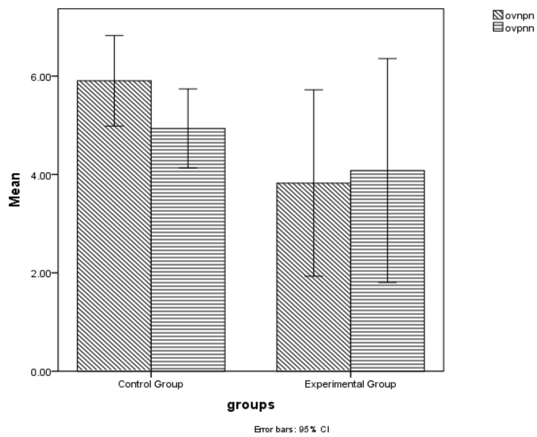


Figure 7: Mean and standard deviation of amplitude (in µV) of n1p1 and p1n2 complexes of ocular VEMP in participants of control and experimental groups.

there was no significant difference between the groups in terms of p1 (Z= -0.574, $p > 0.05$), n1 (Z= -0.034, $p > 0.05$) and n2 (Z= -0.488, $p > 0.05$) latency measures.

Figure 7 depicts the amplitude of n1p1 complex and p1n2 complex for both the control and experimental group. The amplitude of both the n1p1 complex and p1n2 complex was lesser for the experimental group compared to the control group. Mann Whitney U test revealed a statistically significant difference in amplitude of n1p1 complex (Z= -2.862, $p < 0.05$) and also p1n2 complex (Z= -2.478, $p < 0.05$). This implies that the participants with diabetes mellitus had a significantly lesser amplitude for both the cVEMP as well oVEMP. However, there was no significant difference in terms of latencies of the various peaks of cVEMP or oVEMP. Standard deviation in the experimental group was higher compared to the control group for both the n1p1 and p1n2 complex.

Vestibular Symptoms Presented by Individuals with diabetes and their cVEMP and oVEMP Results

The percentage of individuals reporting with each symptom and their cVEMP and oVEMP results were tabulated (Table 3 and Table 4 respectively) for participants in the experimental group.

To summarize the results, the cVEMP and oVEMP could be recorded in all the participants of the control group whereas, cVEMP for the experimental group could be recorded only from 12 ears and oVEMP could be recorded for only 19 ears. Further, there was no correlation between the cVEMP and oVEMP results in both the control as well as the experimental group. Also, there was no correlation between the duration

Table 4: Percentage of individuals reported with various vestibular symptoms and their oVEMP results

	cVEMP right				cVEMP left			
	A	N	Ab	Total %	A	N	Ab	Total %
Light headedness or swimming sensation in the ear	3(50%)	2(66.6%)	1(33.3)	6(20%)	2(33.3%)	2(33.3%)	2(33.3%)	6(20%)
Blacking out or loss of consciousness	1(50%)	1(50%)	0	2(6.7%)	2	0	2(100%)	2(6.7%)
Tendency to fall	9(90%)	1(10%)	0	10(33.3%)	10(100%)	0	0	10(33.3%)
Objects spinning or turning around you	5(71.4%)	0	2(28.5%)	7(23.3%)	3(42.8%)	1(14.2%)	1(42.8%)	7(23.3%)
Sensation that you are turning or spinning inside	1(100%)	0	0	1(3.3%)	1(100%)	0	0	1(3.3%)
Loss of balance when walking	2(50%)	2(50%)	0	4(13.3%)	2(50%)	2(50%)	0	4(13.3%)
Total	21(70 %)	6(20 %)	3(10%)	30	20(60%)	5(16.6%)	5(23.3%)	30

(Note: oVEMP= Ocular Vestibular Evoked Myogenic Potential; A= Absent, N= Normal, Ab= Abnormal)

of the diabetes and puretone average with the cVEMP and oVEMP results. Latency of neither cVEMP nor the oVEMP showed significant differences across the control and experimental groups. Whereas, the amplitude values of both the cVEMP as well as the oVEMP was significantly lesser for the experimental group compared to the control group.

Discussion

Latency and Amplitude of cVEMP and oVEMP in Control Group

All the participants of the control group in the present study had presence of cVEMP and oVEMP. The latency of P1, N1 and amplitude of P1N1 complex of cVEMP and latency of n1, p1 and amplitude of n1p1 complex obtained in the present study was similar to the studies reported in the literature earlier (**Akin & Murnane** 2001; **Akin, Murnane & Medley** 2003; Bohra, Sanju & Sinha, 2012; Colebatch et al., 1994; Chiarovano, Zamith, Vidal & de Waele, 2011; Murnane, Akin, Kelly & Byrd, 2011; **S**mulders et al., 2009; **Welgampola & Colebatch,** 2001).

During the oVEMP recording an additional negative peak n2 was identified which occurred immediately after the p1 peak. The 'n2' peak of oVEMP was present in all the participants of the control group at a latency of 22 to 23 ms. The presence of 'n2' peak in oVEMP has not been reported earlier in any of the studies. It is hypothesized that the generators of the 'n2' peak also might be confined in the same anatomical structures from where the 'n1' and 'p1' peak is generated.

The first peak of the oVEMP in the present study was recorded around a mean latency of 11.97 ms and 12.16 ms for right and left ear respectively, whereas first peak of cVEMP was recorded with a mean latency of 14.89 ms and 14.75 ms for right and left ear respectively. Thus, the latency measures indicate that the latency for the oVEMP is shorter compared to that of cVEMP. The differences in latencies between cVEMPs and oVEMPs might be due to the differences in length and nerve conduction velocity between vestibular ocular (Broussard & Lisberger, 1992) and vestibulo spinal pathways (Uchino et al., 2005) as shown in animal studies. cVEMPs and oVEMPs responses obtained by air conduction stimulation are generated from different anatomical pathways (Chiarovano et al., 2011) and follows a different pathway before the muscle potential (Sternocleidomastoid and Contralateral extraocular muscle) is recorded (Halmagyi & Curthoys, 1999; Rosengren, Welgampola & Colebatch, 2010). The pathway for the oVEMP is relatively shorter compared to that of cVEMP (Rosengren et al., 2010) and hence the latency differences were obtained in the present study.

Also, the amplitude obtained for the P1N1 complex was higher compared to the oVEMP in the present study. The differences in amplitude between cVEMP and oVEMP may be due to the differences in the muscle unit content between SCM and extraocular muscles (Park et al., 2010). The muscle thickness is reported to be more at the SCM compared to the extraoccular muscles and hence the tonic activation is more for the SCM comared to the extraoccular muscles (Park et al., 2010).

cVEMP and oVEMP Results in the Experimental Group

The results indicate a normal functioning of the cochlear nerve since the auditory brainstem responses are normal. Since cochlear and the vestibular nerves are the part of the same 8^{th} cranial nerve, it can hypothesised that the vestibular nerve functions are also normal in diabetic individuals. The primary site of lesion in the individuals with diabetes might be confined to the end organs specifically in the utricle or the saccule.

Myers, Ross, Jokelainen, Graham and McClatchey (1985) have demonstrated vestibular end-organ pathological changes, such as increased capillary diameter of the small blood cells of the utricle and saccule and accumulation of lipid droplets in subneuroepithelial connective tissue cells of these vestibular organs. It is common for capillaries within a capillary bed to vary in size but it is noteworthy that over 25 % of the control capillaries in persons with diabetes were under 4 μm in diameter (Myers et al., 1985). The higher viscosity of diabetic blood (Schmid-Schonbein & Volger, 1976) and the decreased deformability of diabetic red blood cells (McMillan, Utterback & La Puma, 1978; McMillan & Gion, 1981; Otsuji, Baba & Kamada 1981) are likely to cause impaired blood flow through such narrow channels. Under these conditions, either the passage of red blood cells would be slowed down reducing oxygen delivery to the tissues, or the mechanical force exerted on the capillary wall would be increased. The latter situation is considered a strong candidate in the development of diabetic microangiopathy (McMillan, 1983) and could explain the increased capillary diameters in the saccules and utricles of the diabetic subjects.

The increased density of capillaries in the saccules and utricles of the diabetic subjects indicates vascular proliferation induced by the presence of diabetes and that this proliferation occurs within the first three months of diabetes (Myers et al., 1985). A vascular proliferation such as this would be expected to be a reflection of either an increased oxygen demand by the tissue or alternatively, by a decreased efficiency of oxygen delivery by the capillary bed. No evidence is available to support the former possibility. The latter case is supported by studies which have shown that the glycosylation of hemoglobin in diabetic blood increases the oxygen affinity of the hemoglobin thereby impairing the release of oxygen to the tissues (Ditzel, 1976; Bunn, Gabbay & Gallop, 1978).Reduced oxygen delivery by diabetic blood has been challenged (Bunn et al.,1978)however, on the grounds that other physiological variables in the blood which influence oxygen release would negate the affect of the increased oxygen affinity of glycosylated hemoglobin. If this is the case, then the possibility of a reduced flow rate, mentioned earlier with regard to elevated blood viscosity, could be an alternative cause of a relative hypoxia of the saccule and utricle leading to a damage of the saccule and utricle and hence absence of cVEMP and oVEMP.

There was no difference in the latency of P13 peaks or N23 peaks among the control versus experimental groups. Various studies have reported that latency parameter of VEMP is relatively less subject to undergo changes than amplitude and threshold of VEMP response (Faith et al., 2004). Also, the latency parameters were insensitive to stimulus characteristics (Faith et al., 2004).

Other studies which involved the study of degeneration process of the sacculocollic pathways and other pathological conditions have also reported no significant change in the latency parameters compared to the amplitude parameters (Kumar et al., 2007; Murofushi, Matsuzaki & Takegoshi 2001; Sun Kyu Lee, et al., 2007; Welgampola & Colebatch, 2001; Young, Huang & Cheng, 2003). However, in older population a prolongation in the latency of VEMP has been reported (Lee et al., 2008; Kumar, Sinha & Bhat, 2011), indicating that the latency parameters might be sensitive only in the nerve pathology rather than the end organ pathologies. In end organ pathologies the amplitude parameter seem to be more sensitive in detecting the pathology. Thus, no difference in the latency of cVEMP and oVEMP between the persons with and without diabetes mellitus is an indication that pathology could be restricted to the otolith organs of the diabetic subjects and not the vestibular nerves.

The results of the present study also indicate that in most of the participants, both the saccule as well as the utricle was involved bilaterally. The presence of cVEMP in the absence of oVEMP indicates a possible involvement of the utricle alone whereas, the absence of cVEMP in the presence of oVEMP indicates involvement of the saccule alone. Overall, the saccule was more involved compared to the utricle in individuals with diabetes mellitus suggesting a higher susceptibility of saccular end organ to the deoxygenation caused by diabetes compared to the utricle.

Correlation of Latency and Amplitude Measures of cVEMP and oVEMP in Control and Experimental Group

No correlation was obtained for the latency or amplitude measures of cVEMP and oVEMP for control group as well as the experimental group. There is dearth of information regarding the correlation between the cVEMP and oVEMP measures in normal healthy subjects. However, there are studies which have reported a poor correlation between the cVEMP and oVEMP in individuals with Meniere's disease (Chiarovano et al., 2011; Murofushi, Nakahara, Yoshimura & Tsuda, 2011). Murofushi etal., (2011) reported poor correlation of cVEMP and oVEMP, while oVEMP latency and amplitude measures significantly correlated with the caloric test in individuals with Meniere's disease. Similar findings were also reported by Chiarovano et al. (2011) who also concluded that the cVEMP responses are generated majorly in the saccular region while the oVEMP responses are from the utricular region.

Lack of correlation between the cVEMP and oVEMP measures has been attributed to the two different pathways involved in the generation of these potentials

(Huang, Wang & Young, 2012). The cVEMP test runs via the inferior vestibular nerve pathways, whereas the oVEMP runs via the superior vestibular nerve pathways (Huang et al., 2012). Also, since the amplitude of VEMP is dependent upon the muscle tension, no correlation between the cVEMP and oVEMP amplitude might be due to the fact that the muscle thickness for ocular muscle is lesser compared to the sternocleidomastoid muscles (Park et al., 2010). There might not be a significant correlation between the cVEMP and oVEMP but combining test of oVEMPs and cVEMPs may provide localization of pathology (Huang, Wang & Young, 2011).

Correlation between the Duration of Diabetes and Pure Tone Average with cVEMP and oVEMP Results

There was no correlation between the duration of diabetes with cVEMP or oVEMP results. Also, no significant correlation was obtained between puretone average and cVEMP and oVEMP results. Rajendran, Anandhalakshmi, Mythili and Rao (2011) reported that the duration of diabetes (above or below 10 years) has no effect in the incidence of hearing loss in the diabetic group. No correlation was reported between duration of diabetes and degree of hearing loss (Panchu, 2008) and duration of diabetes and abnormality of auditory brainstem responses (Zehra, Kaya, Gonen & Ilhan, 1999). Duration of diabetes might not be a significant factor in abnormality of the different tests rather the uncontrolled levels of the glucose might be a factor in damaging different structures (Panchu, 2008) such as the otolith organs as found in this study.

Most of the participants in the present study had minimal hearing loss (<30dBHL), a finding supported by many studies in literature. Several probable mechanisms of hearing loss in cases with diabetes have been proposed such as microangiopathy of the inner ear, neuropathy of the cochlear nerve, a combination of both, outer hair dysfunction and disruption of endolymphatic potential. The tissue effects of diabetes are thought to be related to the polyol pathway, where glucose is reduced to sorbitol. Sorbitol accumulation is implicated in neuropathy by causing a decrease in myoinositol content, abnormal phosphoinositide metabolism and decrease in Na+ K+ ATPase activity (Dennis et al., 2008). Since the auditory ABRs were normal in present study in all the participants of experimental group, it can be hypothesized that the lesion in individuals who participated for this study might be confined to the cochlear structures. Makishima and Tanaka (1971) have also reported a severe atrophy of the spiral ganglion in the basal and middle turns of the cochlea in diabetic patients with sensorineural hearing loss. Further no correlation between the puretone average and VEMP (cVEMP and oVEMP) results could be due to the fact that the structures involved in processing of the puretone signals and generation of VEMP (cVEMP and oVEMP) are different. It can be hypothesized that the level of glucose might have a differential effect on the two structures i.e it might affect the vestibular structures more than cochlear structures. However, there are no studies to support or refute this hypothesis.

Sign and Symptoms Exhibited by the Individuals with Diabetes and VEMP Results

On administering the Maryland dizziness questionnaire all the participants reported one or the other vestibular symptoms. Further, the client who exhibited these symptoms, most of them had abnormal/absent cVEMP and oVEMP. Based on the findings of this study (i.e absence of both the cVEMP and oVEMP in most of the subjects), it is expected that many more diabetic patients will have vestibular symptoms. Initially, when the general case history was taken, none of the participants reported vestibular symptoms. On intentional screening using the Maryland dizziness questionnaire vestibular symptoms were revealed indicating the importance of additional vestibular screening in the test battery for persons with diabetes mellitus. Most of the clients with vestibular symptoms had absent VEMP responses which indicated an otolith organ disorder (both utricle and saccule). The symptoms exhibited by the individuals with diabetes might be secondary to the damage of the otolith organs.

First, it is possible that this is not seen clinically because vertigo would probably be reported in the presence of functional asymmetry between the two inner ears. In diabetes mellitus and other metabolic diseases it is expected that there is a symmetrical impairment, which probably causes clinical symptoms only when the impairment severity increases. Second, some central compensation might be taking place in these individuals and hence they do not exhibit these symptoms clinically. Therefore, detail information about the vestibular symptoms should be collected from individuals with diabetes.

Conclusions

Diabetes can affect different vestibular structures. The site of lesion in individuals with diabetes can be confined to end organs rather than the neural system. The amplitude measures of cVEMP and oVEMP are more sensitive parameters than the latency measures. As the vestibular system is complex involving multiple structures, one must administer different tests to rule out pathology of each of these structures. Mostly the persons with diabetes remain asymptomatic probably because of bilateral distribution of the disorder or a probable central compensation, hence a detailed case history must be taken to rule out or detect any vestibular symptoms.

The study provides a thrust to long felt need for research in the field of vestibular assessment in individuals with diabetes. Present study opens a new research era in understanding the involvement of the sacculocollic and the utriculooccular pathway in individuals with diabetes. The vestibular evoked myogenic potentials may give an indication of the involvement of the different pathways of the vestibular system. Vestibular rehabilitation therapy (VRT) exercises are typically based on principles of vestibular adaptation of semicircular canal input. If otolith organ involvement is identified, then VRT exercises designed to stimulate otolithic adaptation may be more effective for managing a patient's symptoms.

References

Adriano, S., & Silvia, A. V. (2006). Correlation between dizziness and impaired glucose metabolism. *Brazilian Journal of Otolaryngology, 72*(3), 367-369.

Akin, F. W., & Murnane, O. D. (2001). Vestibular evoked myogenic potentials: A preliminary report. *Journal of the American Academy of Audiology, 12*(9), 445-452.

Akin, F. W., Murnane, O. D., & Medley, T. M. (2003). The effects of click and tone burst stimulus parameters on the vestibular evoked myogenic potential (VEMP). *Journal of the American Academy of Audiology, 14*(9), 500-508.

ANSI (1999). *American National Standard: Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms.* New York: American National Standards Institute, Inc., ANSI S3.1 1999.

Bohra, V., Sanju, H. K., & Sinha, S. K. (2012). *A comparative study of cVEMP and oVEMP between dancers and non-dancers.* Research paper presented at 44th ISHA conference in Hyderabad.

Broussard, D. M. & Lisberger, S. G. (1992). Vestibular inputs to brain stem neurons that participate in motor learning in the primate vestibulo ocular reflex. *Journal of Neurophysiology, 68*(5), 1906-1909.

Bunn, H. F., Gabbay, K. H., & Gallop, P. M. (1978). Glycosylation of hemoglobin relevance to diabetes mellitus. *Science, 200,* 21-27.

Carhart, H., & Jerger, J. F. (1959). Preffered method for clinical determination of pure tone thresholds. *Journal of Speech and Hearing Disorders, 24*(4), 330- 345.

Chiarovano, E., Zamith, F., Vidal, P. P., & deWaele, C. (2011). Ocular and Cervical VEMPs: a study of 74 patients suffering from peripheral vestibular disorders. *Clinical Neurophysiology, 122*(8), 1650-1659.

Colebatch, J., Halmagyi, G., & Skuse, N. (1994). Myogenic potentials generated by a click- evoked vestibulocollic reflex. *Journal of Neurology, Neurosurgery, and Psychiatry, 57*(2), 190-197.

Dennis L., Anthony S., Dan L., Eugene B., Stephen L., Hauser, J. L., & Joseph L. (2008). *Harrison's Principles of Internal Medicine.* McGraw-Hill Companies Inc. United States of America.

Ditzel, J. (1976). Oxygen transport impairment in diabetes. *Diabetes, 25*(2), 832- 838.

Dizziness Questionaire. (2004). Maryland haring and Balance center. Retrieved from http:// free download. is/ doc/ dizziness-questionnaire-1020692.html on 29th September 2011.

Faith A. W., Murnane, O. D., Panus, P. C., Caruthers, S. K., Wilkinson, A. E., & Proffit, T. M. (2004). The influence of voluntary tonic EMG level on the vestibular evoked myogenic potential. *Journal of Rehabilitation Research & Development, 41*(3), 473-480.

Halmagyi, G. M., & Curthoys, I. S. (1999). Clinical Testing of otolith function. *Annals of the New York Academy of Sciences, 871,* 195- 204.

Huang, C. H., Wang, S. J., & Young, Y. H. (2011). Localization and prevalence of hydrops formation in Meniere's disease using a test battery. *Audiology and Neuro otology, 16*(1), 41- 48.

Huang, C. H., Wang, S. J., & Young, Y. H. (2012). Correlation between caloric and ocular vestibular evoked myogenic potential test results. *Acta Otolaryngologica, 132*(2), 160-166.

Jadzinsky, M., Faerman, I., & Fox, D. (1973). Visceral diabetic neuropathy. *Acta Diabetologica latina, 10*(2), 208-260.

Jauregui-Renoud, K., Sanchez, B., Olmos, I. A., & Gonzalez- Barcena, D. (2009). Neuro-otologic symptoms in patients with type 2 diabetes mellitus. *Diabetes Research and Clinical Practice. 84*(3), 45-47.

Kumar, K., Sinha, S. K., Bharti, A. K., Singh, N. K., & Barman, A. (2007). Vestibular evoked myogenic potentials as a tool to identify vestibular involvement in auditory neuropathy. *Asia Pacific Journal of Speech Language and Hearing, 10,* 181-187.

Kumar, K., Sinha, S.K., & Bhat, J.S. (2010). Effect of Aging on Vestibular Evoked Myogenic Potential.*Journal of Indian Speech and Hearing Association, 24*(2), 95-99.

Lee, S. K., Cha, C. I., Jung, T. S., Park, D. C., Yeo, S.G. (2008). Age-related differences in parameters of vestibular evoked myogenic potentials. *Acta Otolaryngologica, 128*, 66-72.

Maia, C. A. S., & Campos, C. A. H. (2005). Diabetes mellitus as etiological factor of hearing loss. *Revised Brazilian Journal of Otorhinolaryngology, 71,* 208-214.

Makishima, K., & Tanaka, K.(1971). Pathological changes of the inner ear and central auditory pathways in diabetics. *Annals of Otology Rhinology and Laryngology, 80,* 218-288.

McMillan, D. E. (1983). The effect of diabetes on blood flow properties. *Diabetes, 32*(2), 56-63.

McMillan, D. E., & Gion, K. M. (1981). Glucosylated hemoglobin and reduced erythrocyte deformability in diabetes, *Hormone and Metabolic Research Supplement, 11,* 108-112.

McMillan, D. E., Utterback, N. G., & La Puma, J. (1978). Reduced erythrocyte deformability in diabetes, *Diabetes, 27*(9), 895-901.

Murnane, O. D., Akin, F. W., Kelly, K. J., & Byrd, S. (2011). Effects of stimulus and recording parameters on the air conduction ocular vestibular evoked myogenic potential. *Journal of the American Academy of Audiology, 22*(7), 469-480.

Murofushi, T., Matsuzaki, M., & Takegoshi, H. (2001). Glycerol affects vestibular evoked myogenic potentials in Meniere's disease, *Ear Nose and Throat, 28*, 205-208.

Murofushi, T., Nakahara, H., Yoshimura, E., & Tsuda, Y. (2011). Association of air conducted sound oVEMP findings with cVEMP and caloric test findings in patients with unilateral peripheral vestibular disorders. *Acta Otolaryngologica, 131*(9), 945-950.

Myers, S. F., Ross, M. D., Jokelainen, P., Graham, M. D., & McClatchey, K. D. (1985). Morphological evidence of vestibular pathology in long term experimental diabetes mellitus: I. Microvascular Changes. *Acta Otolaryngologica, 100*(5-6), 351-364.

Otsuji, S., Baba, Y., & Kamada, T. (1981). Erythrocyte membrane microviscosity in diabetes. *Hormone and Metabolic Research Supplement Series,11,* 97-102.

Panchu, P. (2008). Auditory acuity in type 2 diabetes mellitus. *International Journal of Diabetes in Developing Countries, 28*(4), 114-120.

Park, H., Lee, I., Shin, J., Lee, Y., & Park, M. (2010). Frequency tuning characteristics of cervical and ocular vestibular evoked myogenic potentials induced by air conducted tone bursts. *Clinical Neurophysiology, 121*, 85-89.

Petrofsky ,J., Lee, S., Macnider, M., & Navarro, E. (2005). Autonomic, endothelial function and the analysis of gait in patients with type 1 and type 2 diabetes. *Acta Diabetologica, 42*(1), 7-15.

Petrofsky, J. S., Cuneo, M., Lee, S., Johnson, E., & Lohman, E. (2006). Correlation between gait and balance in people with and without Type 2 diabetes in normal and subdued light. *Medical Science Monitor, 12*(7), 273-281.

Petrofsky, J., Lee, S., & Bweir, S. (2005). Gait characteristics in people with type 2 diabetes mellitus. *European Journal of Applied Physiology, 93*(5-6), 640- 647.

Rajendran, S., Anandhalakshmi., Mythili, B., Viswanatha, Rao. (2011). Evaluation of the incidence of sensorineural hearing loss in patients with type 2 diabetes mellitus. *International Journal of Biological and Medical Research, 2*(4): 982-987.

Rosengren, S. M.,Welgampola, M. S., & Colebatch, J. G. (2010). Vestibular evoked myogenic potentials: past present and future. *Clinical Neurophysiology, 121*(5), 636-651.

Schmid- Schonbein, H., & Volger, E. (1976). Red cell aggregation and red cell deformability in diabetes, *Diabetes, 25*(2), 897-902.

Shagan, B. P. (1976). Diabetes in the elderly patient. *The Medical Clinics of North America, 60*(6), 1191-1208.

Smulders, Y. E., Welgampola, M. S., Burgess, A. M., McGarvie, L. A., Halmagyi, G. M., & Curthoys, I. S. (2009).The n10 component of the ocular vestibular-evoked myogenic potential is distinct from the R1 component of the blink reflex. Clinical Neurophysiology, *120*, 1567-1576.

Sun, K. L., Chang, C., Tae, S. J., Dong C, P., & Seung, G. Y. (2007). Age-related differences in parameters of vestibular evoked myogenic potentials. *Acta Otolaryngologica, 128(1)*, 66-72.

Tavormina, J. G., Kastner, L. S., Slater, P. M., & Watts, S. L. (1976). Chornologically ill children, A psychologically and emotionally deviant population, *Journal of Abnormal Child Psychology, 4*(2), *99-110.*

Todd, N., Rosengren, S., & Colebatch, J. (2003). A short latency vestibular evoked potential (VsEP) produced by bone conducted acoustic stimulation. *Journal of the Acoustical Society of America, 114*(6), 3264- 3272.

Uchino, Y., Sasaki, M., Sato, H., Bai, R., & Kawamoto, E. (2005). Otolith and canal integration on single vestibular neurons in cats. *Experimental Brain Research, 164*, 271-285.

Welgampola, M. S., & Colebatch, J. G. (2001). Characteristics of tone burst evoked myogenic potentials in the sternocleidomastoid muscles. *Otology and Neurootology, 22*(6), 796-802.

Young, Y. H., Huang, T. W., & Cheng, P. W. (2003). Assessing the stage of Meniere's disease using vestibular evoked myogenic potentials. *Archives of Otolaryngology and Head and Neck Surgery, 129,* 815-818.

Zehra, A., Kaya, A., Gonen, S., & Ilhan, N. (1999). Brainstem auditory evoked potentials in patients with type-2 diabetes mellitus. *Turkish Journal of Endocrinology and Metabolism, 1,* 29-32.

Brainstem Correlates of Speech Perception in Noise: Carnatic Musicians Vs. Non-Musicians

[1]Zubin Vinod & [2] Rajalakshmi K.

Abstract

Many studies have indicated the presence of superior auditory capabilities as a result of long-term musical experience, including better perception of speech in a background of noise. Musicians have life long experience parsing melodies from background harmonies, which can be considered a process analogous to speech perception in noise. To investigate the effect of musical experience on the neural representation of speech-in-noise, the subcortical neurophysiological responses to speech in quiet and noise in a group of highly trained musicians and nonmusician controls were compared. Musicians were found to have a more robust subcortical representation of the acoustic stimulus in the presence of noise. Specifically, musicians demonstrated earlier latencies and higher amplitudes of onset and transition peaks, higher amplitudes of encoded formants and less degraded response morphology in noise. Neural measures were associated with better behavioral performance on the test of Speech Perception in Noise (SPIN) for which musicians outperformed the nonmusician controls. These findings suggest that musical experience limits the negative effects of competing background noise, thereby providing the first biological evidence for musicians' perceptual advantage for speech-in-noise.

Keywords: *Music, Carnatic, Speech ABR, FFR, SPIN*

Introduction

The domains of music and language share many features, the most direct being that both exploit changes in pitch patterns to convey information. Music uses pitch contours and intervals to communicate melodies and tone centers. Pitch patterns in speech convey prosodic information; listeners use prosodic cues to identify indexical information, i.e., information about the speaker's intention as well as emotion and other social factors.

Through years of sensory-motor training, often beginning in early childhood, musicians develop an expertise in their instrument of specialization or mastery over their voice. In the course of training, musicians increasingly learn to attend to the fine-grained acoustics of musical sounds. Attention to pitch, timing and timbre is emphasized during music training. A variety of studies have found that musical training improves auditory-perceptual skills resulting in enhanced behavioural (Jeon & Fricke, 1997; Koelsch, Schroger & Tervaniemi, 1999; Micheyl Delhommeau, Perrot & Oxenham, 2006; Rammsayer & Altenmuller, 2006; Tervaniemi, et al., 2009) and neurophysiological (Brattico, Naatanen & Tervaniemi, 2001; Pantev et al., 2001; Schneider, et al., 2002; Shahin , Bosnyak , Trainor, Roberts & Larrey, 2003; Tervaniemi, et al., 2005; Kuriki, Kanda, & Hirata, 2006; Kraus, Skoe, Parbery-Clark & Ashley, 2009) responses.

It is only reasonable to assume that the benefits that musicians have in processing music would also extend to speech stimuli. A number of research studies have shown that music training benefits auditory processing not only in the musical domain, but also in the processing of speech stimuli (Musacchia et al., 2007; Schon, Magne & Besson, 2004; Wong, Skoe, Russo, Dees & Kraus, 2007). Other verbal and non-verbal skills such as auditory attention (Strait, Kraus, Parbery-Clark, & Ashley, 2010), auditory stream segregation (Beauvois & Meddis, 1997), processing emotion in speech (Strait, Kraus, Skoe & Ashley, 2009), working memory (Chan, Ho & Cheung, 1998; Forgeard, Winner, Norton & Schlaug, 2008) and processing of prosody and linguistic features in speech (Chandrasekaran, Krishnan & Gandour, 2009; Wong, Skoe, Russo, Dees & Kraus, 2007).

Of special note is the enhanced ability of musicians to extract relevant signals from a complex soundscape (e.g., the sound of their own instrument in an orchestra). Speech perception in noise is a complex task that requires the segregation of target signals from a competing background noise. To complicate matters, the noise also degrades the signal particularly by disrupting the perception of rapid spectro-temporal changes (Brandt & Rosen, 1980). Poor performance in the task of speech perception in noise is seen in individuals with hearing impairment (Gordon- Salant & Fitzgibbons, 2004) and language-based learning disabilities (Bradlow, Kraus & Hayes, 2003; Ziegler, Pech-Georgel, George & Lorenzi, 2005) whereas musicians demonstrate better performance than non-musicians (Parbery-Clark, Skoe & Kraus, 2009). It was hypothesized that a musician's long-term experience with musical stream segregation would transfer to the homologous task of speech perception in noise. Parbery-Clark et al. (2009) found a distinct speech in noise advantage for musicians, as measured by two standardized

[1]Email: zforce432@gmail.com,
[2]professor of Audiology, Email: veenasrijaya@gmail.com

tests of hearing in noise (HINT, Hearing in-noise test; QuickSIN). Musicians showed superior working memory and performed better on a frequency discrimination task. Across all participants, the number of years of consistent practice with a musical instrument correlated strongly with performance on QuickSIN, auditory working memory and frequency discrimination. These correlations strongly suggest that practice fine tunes cognitive and sensory ability, leading to an overall advantage in speech perception in noise in musicians.

All these enhanced abilities in musicians may be related to structural and functional enhancements seen at different levels of their nervous system. For instance, musicians have more neural cell bodies (grey matter volume) in the auditory, motor and visuo-spatial areas of the brain (Gaser & Schlaug, 2003) and also have more axonal projections that connect the right and left hemispheres (Schlaug, Jancke, Huang, Staiger & Steinmetz, 1995). All these anatomical enhancements are seen to translate into improved auditory and cognitive skills as is evidenced by various studies. The intensive practice over the years has been attributed to bring about neuroplastic changes in the practitioner as is evidenced in many research studies (Pantev et al., 1998; Koelsch et al., 1999; Pantev, Roberts, Schulz, Engelien & Ross, 2001; Tervaniemi, Rytkonen, Schroger, Ilmoniemi & Naatanen, 2001; Fujioka, Trainor, Ross, Kakigi & Pantev, 2005; Musacchia, Sams, Skoe & Kraus, 2007;). One of the mechanisms used to explain the findings of music-induced experience dependent plasticity at the level of the brainstem is increased efficiency of top-down predictive coding (Strait et al., 2010). Recent studies have suggested an important role for the feedback (top-down) pathways in fine-tuning the auditory signal at early stages of auditory processing (Luo, Wang, Kashani & Yan, 2008). Such top-down influences back-project all the way to the cochlea through the medial olivocochlear bundle (MOCB). These authors have said that feedback initiated by the higher (cortical) structures is transferred to the lower (brainstem) structures via the efferent auditory system. This results in an enhanced selectivity of sound features at the lowest levels of the auditory system which is important for higher-level structures to distinguish relevant information in the signal from irrelevant details. The human auditory brainstem response (ABR) has been used as an index of brainstem encoding of speech stimuli (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010).

Since the FFR preserves spectral information up to about $\tilde{2}000$ Hz and reflects neural timing in the order of milliseconds, it can therefore be used to examine the fidelity of the brainstem representation of spectral and timing information. It has been found that the addition of background noise delays the timing of brainstem responses (Cunningham, Nicol, Zecker & Kraus, 2000; Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001) and reduces spectral magnitude. There is evidence from studies using speech-evoked ABR that music training modulates the effect of background noise on subcortical auditory representation (Parbery- Clark, et al., 2009). Musicians show less degraded brainstem representation of speech relative to non-musicians, as evidenced by faster neural timing, enhanced spectral representation, and better stimulus-to-response correlations. Though the differences between musicians and non-musicians are present even in quiet backgrounds (Musacchia et al., 2007), it is in the presence of background noise that the differences in spectral representation between musicians and non-musicians are large, suggesting that musical experience protects against the debilitating effects of background noise (Parbery-Clark, et al., 2009). Thus timing and spectral features are preserved at the level of the brainstem to a greater extent due to musical experience and these enhancements translate into a better performance on the task of speech perception in noise. The speech-evoked ABR is hence considered to be a reliable indicator of the biological basis of speech perception in noise.

Despite the considerable amount of literature dealing with the enhanced subcortical encoding of speech in the presence of noise in Western musicians, there is a dearth of similar studies in Carnatic musicians. Thus the following study was carried out to verify whether trained Carnatic musicians show better perception of speech in the presence of background noise as compared to non-musicians and if so, whether they had enhanced subcortical encoding of speech stimuli as measured via speech evoked ABR as compared to non-musicans.

Method

Subjects

Fifteen musicians and fifteen nonmusicians participated in this study. Participants' age ranged from 18 to 30 years. Participants categorized as musicians started training in Carnatic music before the age of 8 and practiced consistently for at least10 years before enrolling in the study. Nonmusicians were required to have had no musical training. All participants had normal hearing thresholds from 125 to 8000 Hz. No participant reported any cognitive or neurological deficits.

Stimuli

The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980). This stimulus simultaneously contains the broad spectral and fast temporal information characteristic of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. Although the steady-state portion is not present, the stimulus is still perceived as being a consonant-vowel syllable. The fundamental frequency (F0) linearly rises from

103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 ms. The first formant (F1) rises from 220 to 720 Hz, while the second formant (F2) decreases from 1700 to 1240 Hz over the duration of the stimulus. The third formant (F3) falls slightly from 2580 to 2500 Hz, while the fourth (F4) and fifth formants (F5) remain constant at 3600 and 4500 Hz, respectively. The phonemically balanced wordlist in Kannada (Yathiraj & Vijayalakshmi, 2005) was presented in the presence of ipsilateral speech noise to assess the patient's perception of speech in noise.

Procedure

The speech syllable /da/ was presented in condensation and rarefaction polarities at 80 dB sound pressure level (SPL) through insert ear phones (ER-3; Etymotic Research). In the noise condition, both the /da/ and white noise were presented simultaneously to the test ear. The /da/ was presented at a 0 dB signal-to-noise ratio over the background noise.

The responses to two background conditions, quiet and noise, were collected using Bio-Logic Navigator Pro EP with 3 gold disc electrodes which were fastened to the scalp. Responses were differentially recorded with a vertical montage (Cz active, forehead ground, and earlobe references), an optimal montage for recording brainstem activity (Galbraith et al., 1995; Chandrasekaran & Kraus, 2009). Contact impedance was 2 kΩ or less between electrodes. Three thousand artifact-free sweeps were recorded for each condition for both polarities. Participants were asked to sleep for the recording session. To limit the inclusion of low-frequency cortical activity, brainstem responses were off-line bandpass filtered from 70 to 2000 Hz (12 dB/octave, zero phase-shift) using Bio-logic Navigator Pro EP. The filtered recordings were epoched using a time window of 64 ms which included a prestimulus time of 10 ms (default setting in Biologic system) with the stimulus onset occurring at 0 ms. Any sweep with activity greater than 35μV was considered artifact and rejected. The responses to the two polarities were added together to minimize the presence of the cochlear microphonic and stimulus artifact on the neural response (Gorga et al., 1985; Aiken & Picton, 2008). Last, responses were amplitude-baselined to the prestimulus period.

Analysis

The latency and amplitude of onset peak V and transition peaks D, E and F were measured. The waveforms obtained in both conditions were also Fast Fourier Transform (FFT) to obtain information regarding the spectral characteristics of the FFR (frequency and amplitude of spectral peaks). The average spectral amplitude was calculated for a frequency range from 103 to 120 Hz which encompasses the fundamental frequency (F0). FFT was performed on all speech evoked potentials using a custom made program run in MATLAB. The peak amplitude corresponding to F0 was also calculated using a custom made program file in the MATLAB platform. The frequency analysis was done from 11.4 to 40.6 ms. The sustained portion of the response (FFR) was passed through 103 to 120Hz band pass fourth order Butterworth filters in order to obtain the energy at F0. The Fourier analysis was then performed on the filtered signal. A subject's responses were required to be above the noise floor in order to be included in the analysis. This was performed by comparing the spectral magnitude of pre stimulus period to that of the response. If the quotient of the magnitude of F0 frequency component of FFR divided by the pre stimulus period was >1, the response was deemed to be above the noise floor. Statistical analyses were done using SPSS 20.

Results

In quiet, the onset peaks V and the transition peaks D, E, and F were clearly visible in the speech evoked ABR of the non-musicians. The morphology of the waves was noticeably poorer in noise, with peaks having reduced amplitude and delayed latencies. As in Figure 1, the V-A complex is almost eliminated in noise, though the transition waves are less affected. In quiet, the morphology of the Speech-Evoked ABR of musicians did not vary much from that seen in non-musicians.

Though the waveform morphology was poorer in noise than in quiet, the waves were by and large better defined than in the corresponding waveforms of non-musicians. The V-A complex in particular is more clearly seen (Figure 2).

Comparison of Peak Latencies

The comparison of latency measures obtained in different conditions across the groups using mixed ANOVA reveals the presence of main effects of conditions [$F_{(1,28)} = 115.146$, $p<0.001$] and groups[$F_{(1,28)} = 27.664$, $p<0.001$] as well as interaction effects between conditions and groups[$F_{(1,28)} = 10.745$, $p=0.003$], latency and groups[$F_{(3,84)} = 20019.337$, $p<0.001$], conditions and latency [$F_{(3,84)} = 9.087$, $p= 0.022$] and conditions,

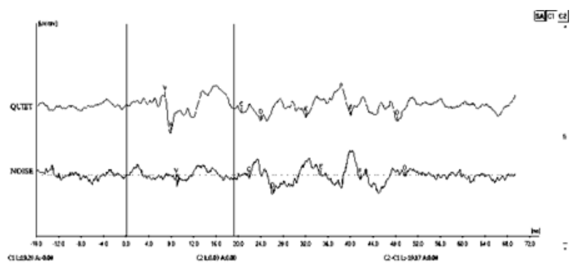


Figure 1: Speech-Evoked ABR in response to 40 ms /da/ acquired in a Non-Musician in quiet and in noise (0dB SNR).

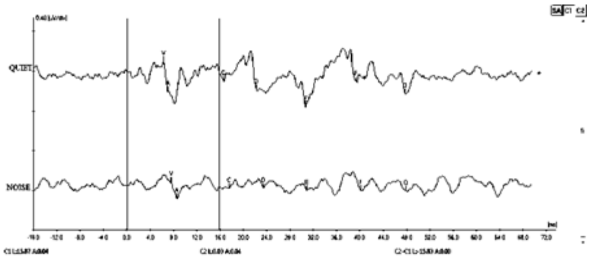


Figure 2: Speech-Evoked ABR in response to 40 m /da/ acquired in a Musician in quiet and in noise (0dB SNR).

latencies and groups [$F_{(3,84)}$ = 3.389, p=0.005].

Descriptive statistics were also done to find out the mean and standard deviation of the latencies for musicians and non-musicians in quiet and in noise. Paired t-test was also carried out to check for the presence of significant differences in the latencies of the waves acquired in quiet and noise in each group.

It is evident from the results that musicians showed earlier mean latencies of all the waves than non-musicians in quiet (Table 1) and in noise (Table 2).

There are also significant differences in the latencies of wave V [$t_{(14)}$= -9.909, p<0.001], wave D [$t_{(14)}$= -6.633, p<0.001], wave E [$t_{(14)}$= -6.859, p<0.001] and Wave F [$t_{(14)}$= -5.135, p<0.001]in quiet and in noise in non-musicians. The latencies of the waves in quiet and in noise were also found to be significantly different for musicians for wave V [$t_{(14)}$= -8.006, p<0.001], wave D [$t_{(14)}$= -3.938, p<0.001], wave E [$t_{(14)}$= -5.121, p<0.001] and wave F [$t_{(14)}$= -7.371, p<0.001].

Table 1: Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Quiet

Measure	Group	**Mean**	Standard Deviation
Wave V Latency	Nonmusicians	6.23	0.32
	Musicians	6.02	0.18
Wave D Latency	Nonmusicians	22.88	0.72
	Musicians	22.01	0.34
Wave E Latency	Nonmusicians	31.04	0.61
	Musicians	30.42	0.32
Wave F Latency	Nonmusicians	39.54	0.66
	Musicians	38.97	0.23

Table 2: Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Noise

Measure	Group	Mean	Standard Deviation
Wave V Latency	Nonmusicians	7.68	0.71
	Musicians	7.43	0.74
Wave D Latency	Nonmusicians	25.35	1.50
	Musicians	22.96	0.91
Wave E Latency	Nonmusicians	33.80	1.65
	Musicians	31.97	1.08
Wave F Latency	Nonmusicians	41.96	2.05
	Musicians	39.89	0.63

Further, one way MANOVA tests were carried out to compare how the latencies of Waves V, D, E and F varied across the groups in quiet and in noise.

The results of the one way MANOVA show that in quiet, the latencies of waveV [$F_{(1,28)}$ = 4.725, p=0.038], wave D [$F_{(1,28)}$ = 17.535, p<0.001], Wave E[$F_{(1,28)}$ = 12.165, p=0.002] and Wave F[$F_{(1,28)}$ = 9.684, p=0.004] were found to be significantly different across the groups. In the presence of noise, the latency of wave V [$F_{(1,28)}$ = 0.890, p=0.353] did not differ significantly across the groups but the latencies of the D [$F_{(1,28)}$ = 27.614, p<0.001], E [$F_{(1,28)}$ = 12.774, p=0.001] and F [$F_{(1,28)}$ = 13.979, p=0.001] did vary.

Comparison of Amplitude Measures

Mixed ANOVA was carried out to compare between the groups for amplitude measures obtained the conditions of quiet and in noise (Table 8). The results of the test reveals the presence of main effects of conditions [$F_{(1,28)}$ = 576.733, p<0.001] and groups[$F_{(1,28)}$ = 19.332, p<0.001] as well as interaction effects between conditions and groups[$F_{(1,28)}$ = 14.248, p=0.001] , amplitude and groups[$F_{(3,84)}$ = 24.940, p<0.001], conditions and latency [$F_{(3,84)}$ = 9.969, p<0.001]and conditions, latencies and groups [$F_{(3,84)}$ = 53.356, p<0.001].

Descriptive statistics were done to find out the mean and standard deviation of the latency measures for musicians and non-musicians in quiet and in noise. The examination of the mean amplitudes of the waves V, D, E and F reveals that the musicians had higher mean amplitudes than non- musicians for all the waves in quiet (Table 3). The mean amplitudes of all the waves acquired in noise were also greater for musicians than for non-musicians (Table 4).

Paired t-test was also carried out to compare the amplitudes of the waves acquired in quiet and noise in each group. In non-musicians, the amplitudes of wave V [$t_{(14)}$= 10.505, p<0.001], wave D [$t_{(14)}$= 5.922, p<0.001], wave E [$t_{(14)}$= 7.388, p<0.001] and wave F [$t_{(14)}$= 7.542, p<0.001] were found to be significantly greater in quiet than in noise. The amplitudes of wave V [$t_{(14)}$= 14.086, p<0.001], wave D [$t_{(14)}$= 15.282, p<0.001], wave E [$t_{(14)}$= 14.826, p<0.001] and wave F [$t_{(14)}$= 3.967, p<0.001] were also found to be significantly greater in quiet than in noise in musicians.

To compare the amplitude measures of the different waves across the 2 groups in quiet and in noise, two

Table 3: Mean and Standard Deviation for Peak Amplitudes (in micro Volts) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean (μV)	Standard Deviation
Amplitude of peak V	Nonmusicians	0.22	0.049
	Musicians	0.27	0.050
Amplitude of peak D	Nonmusicians	0.15	0.073
	Musicians	0.18	0.034
Amplitude of peak E	Nonmusicians	0.22	0.091
	Musicians	0.41	0.097
Amplitude of peak F	Nonmusicians	0.16	0.049
	Musicians	0.17	0.072

Table 4: Mean and Standard Deviation for Peak Amplitudes (in micro volts) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean (μV)	Standard Deviation
Wave V Amplitude	Nonmusicians	0.043	0.031
	Musicians	0.078	0.023
Wave D Amplitude	Nonmusicians	0.050	0.025
	Musicians	0.067	0.015
Wave E Amplitude	Nonmusicians	0.044	0.027
	Musicians	0.056	0.014
Wave F Amplitude	Nonmusicians	0.087	0.053
	Musicians	0.105	0.035

measures of one-way MANOVA were carried out. The results show that in quiet, the amplitudes of the waves V [$F_{(1,28)}$ = 8.196, p=0.008]and E [$F_{(1,28)}$ = 29.932, p=<0.001] are significantly greater in musicians than in non-musicians but the amplitudes of wave D [$F_{(1,28)}$ = 2.130, p=0.156] and wave F [$F_{(1,28)}$ = 0.330, p=0.570] were not. In noise, the amplitudes of waves V [$F_{(1,28)}$ = 12.078, p=0.002] and D [$F_{(1,28)}$ = 4.709, p=0.039] were found to be significantly greater in musicians than in non-musicians. However the differences n amplitudes was not the significant for wave E [$F_{(1,28)}$ = 2.461, p=0.128] and wave F [$F_{(1,28)}$ = 1.172, p=0.288].

Comparison of Formant Amplitude Measures

Mixed ANOVA was carried out to compare between the groups for formant amplitude measures obtained the conditions of quiet and in noise. The results indicated as to the presence of a main effect of condition [$F_{(1,28)}$ = 0.330, p<0.001] while no main effect of group [$F_{(1,28)}$ = 3.328, p=.079] was noted. Interaction effects were present between formant amplitudes & groups [$F_{(2,56)}$ = 3.723, p=0.030] as well as conditions & formant amplitudes [$F_{(2,56)}$ = 64.899, p<0.001] while no significant effects of condition and group [$F_{(1,28)}$ = 0.017, p=0.898] & condition, formant amplitude and group [$F_{(2,56)}$ = 0.061, p=0.941] were noticed.

Descriptive statistics were carried out to find out the mean and standard deviation of the formant amplitude measures for musicians and non-musicians in quiet and in noise. The mean values of the amplitudes of formants F0, F1 and F2 were found to be greater for musicians than non-musicians in quiet (Table 5).The mean formant amplitudes were higher for musicians than non-musicians in waveforms acquired in noise (Table 6).

Paired t-test was also carried out to compare the amplitudes of the formants in quiet and noise in each group. In non-musicians, the amplitudes of F0 [$t_{(14)}$= 6.943, p<0.001], F1 [$t_{(14)}$= 6.872, p<0.001] and F2 [$t_{(14)}$= 0.684, p<0.001] were seen to be significantly greater in quiet than in noise. In musicians, the amplitudes of F0 [$t_{(14)}$= 5.386, p<0.001], F1 [$t_{(14)}$= 5.111, p<0.001] and F2 [$t_{(14)}$= 1.128, p<0.001] were seen to be significantly greater in quiet than in noise. The mean values of the formant amplitudes were seen to be lesser in noise than in quiet.

To compare the amplitude measures of the different waves across the 2 groups in quiet and in noise, two measures of one-way MANOVA were carried out.

The results of the One Way MANOVA tests show that the formant amplitudes of F0 [$F_{(1,28)}$ = 2.401, p=0.132], F1 [$F_{(1,28)}$ = 0.141, p=0.710] and F2 [$F_{(1,28)}$ = 0.104, p= 0.750] were not found to be significantly greater for musician as compared to non musicians and in quiet. In noise, the same trend was observed across both the groups for amplitudes of F0 [$F_{(1,28)}$ = 2.636, p=0.116], F1 [$F_{(1,28)}$ = 1.477, p=0.234] and F2 [$F_{(1,28)}$ = 0.013, p=0.908].

Table 5: Mean and Standard Deviation for Formant Amplitudes (in dB) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean(dB)	Standard Deviation
Amplitude of F0	Nonmusicians	5.10	1.64
	Musicians	6.16	2.05
Amplitude of F1	Nonmusicians	0.62	0.20
	Musicians	0.65	0.24
Amplitude of F2	Nonmusicians	0.19	0.049
	Musicians	0.19	0.063

Table 6: Mean and Standard Deviation for Formant Amplitudes (in dB) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean (dB)	Standard Deviation
Amplitude of F0	Nonmusicians	2.23	1.31
	Musicians	3.14	1.73
Amplitude of F1	Nonmusicians	0.25	0.11
	Musicians	0.33	0.23
Amplitude of F2	Nonmusicians	0.17	0.08
	Musicians	0.17	0.05

Table 7: Results of Independent Sample t-Test for comparing SPIN Scores across the 2 groups

Parameter	| t |	Degrees of Freedom	p(2-tailed)
SPIN scores	-3.500	28	0.002

**The mean difference is significant at the .05 level.*

Table 8: Mean and Standard Deviation of SPIN Score in Musicians and Non-Musicians

Group	Mean	Standard Deviation
Non- Musicians	78.667	3.266
Musicians	82.400	2.529

Comparison of SPIN Scores

The performance of the subjects on the task of speech perception in noise was measured in terms of percentage correct scores on the SPIN test which used the Phonemically Balanced Wordlist in Kannada (Yathiraj & Vijayalakshmi, 2005) presented at 0dB SNR in a background of speech noise. It was speculated that the disruption of neural timing and encoding of stimulus features in the presence of competing noise would be lesser in musicians than in non-musicians, resulting in enhanced performance on the task of speech perception in noise.

Table 9: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Non-Musicians

Parameter	Pearson Correlation	p(2-tailed)
Latency Wave V	-0.478	0.072
Latency Wave D	0.000	0.998
Latency Wave E	-0.084	0.767
Latency Wave F	0.201	0.471
Amplitude of Wave V	-0.170	0.545
Amplitude of Wave D	-0.241	0.388
Amplitude of Wave E	0.334	0.224
Amplitude of Wave F	-0.036	0.900
Amplitude of F0	-0.062	0.826
Amplitude of F1	-0.029	0.920
Amplitude of F2	-0.013	0.964

**The mean difference is significant at the 0.05 level.*

An independent sample t-test was carried out to compare SPIN scores across the 2 groups. The results indicate that the scores differ significantly across the 2 groups (Table 7). It may be seen from Table 8 that the musicians had a higher mean score on the SPIN test than the non-musicians.

To investigate whether the superior performance of musicians over non-musicians on the task of speech perception in noise was related to the differences in the subcortical encoding of speech stimuli across the two groups, Pearson Correlation Coefficient was calculated to check whether the SPIN scores correlated to the different latency (latencies of waves V, D, E and F), amplitude (amplitudes of waves V, D, E and F) and formant amplitude (formant amplitudes of F0, F1 and F2) measures in quiet and in noise.

In non-musicians, the SPIN scores did not correlate with any of the measures obtained in quiet (Table 9). In musicians, the SPIN scores were found to correlate negatively with the latencies of waves D and F obtained in quiet (Table 10).

In noise, it was seen that for non-musicians, the SPIN scores negatively correlated with the latency of wave V and positively correlated with the amplitudes of wave V and D (Table 11).

Thus, poorer performance on the SPIN test was found to be related to the prolongation of onset latency and the reduction of amplitudes of the onset wave V and transition wave D in non-musicians, indicating that addition of noise had resulted in disruption of brainstem timing and a reduction in the amplitude of the responses encoding stimulus features (onset and transition), which had resulted in reduced SPIN scores.

In noise, the SPIN scores of musicians correlated neg-

Table 10: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Musicians

Parameter	Pearson Correlation	p (2-tailed)
Latency Wave V	-0.493	0.062
Latency Wave D	-0.655	0.008
Latency Wave E	-0.334	0.224
Latency Wave F	-0.610	0.016
Amplitude Wave V	0.231	0.408
Amplitude Wave D	-0.239	0.391
Amplitude Wave E	0.451	0.092
Amplitude Wave F	0.158	0.573
Amplitude F0	0.195	0.487
Amplitude F1	0.130	0.644
Amplitude F2	-0.250	0.369

**The mean difference is significant at the 0.05 level.*

Table 11: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Non-Musicians

Parameter	Pearson Correlation	p(2-tailed)
Latency Wave V	-0.788	<0.001
Latency Wave D	0.070	0.804
Latency of Wave E	0.015	0.959
Latency of Wave F	0.296	0.284
Amplitude of Wave V	0.541	0.037
Amplitude of Wave D	0.561	0.030
Amplitude of Wave E	-0.159	0.571
Amplitude of Wave F	0.027	0.924
Amplitude of F0	0.164	0.559
Amplitude of F1	0.284	0.304
Amplitude of F2	-0.081	0.775

**The mean difference is significant at the 0.05 level.*

Table 12: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Musicians

Parameter	Pearson Correlation	p(2-tailed)
Latency Wave V	-0.886	<0.001
Latency Wave D	-0.388	0.153
Latency Wave E	-0.096	0.734
Latency Wave F	-0.692	0.004
Amplitude Wave V	0.010	0.973
Amplitude Wave D	0.250	0.368
Amplitude Wave E	0.453	0.090
Amplitude Wave F	0.006	0.982
Amplitude F0	0.517	0.048
Amplitude F1	0.378	0.165
Amplitude F2	0.338	0.217

**The mean difference is significant at the 0.05 level.*

atively with the latencies of wave V and F (Table 12), indicating that subjects with earlier wave V and F latencies showed better performance on the task of speech perception in noise. Positive correlation was seen with the formant amplitude of F0 obtained in noise, indicating that the superior encoding of F0 in musicians had resulted in enhanced SPIN scores.

Discussion

Wave Morphology

It was seen that the addition of noise to the speech stimulus caused the morphology to deteriorate significantly from the quiet condition in both musicians and non-musicians. Similar findings were reported in studies by Russo, Nicol, Zecker, Hayes and Kraus (2004) and Russo, Nicol, Musacchia and Kraus (2004). However, musicians were seen to have a comparatively better morphology of the waveform in the presence of noise than non-musicians. This is in line with the findings of Parbery-Clark, Skoe and Kraus (2009).

Latencies of Onset and Transition Peaks

The latencies of the onset peak V and the transition peaks D, E and F were considered for analysis. The latencies of the peaks are related to the timing of the features of the stimulus (the onset and transition portions). The addition of background noise had been documented to result in delays in latencies of the peaks of ABR, indicating a disruption in timing of brainstem activity (Don & Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). It has been hypothesized that the disruptive effects of noise on the representation of stimulus features may be limited by long-term musical training which can bring about enhancements of stimulus features at the sub cortical level via top down influences (Dean, Harper & McAlpine, 2005) mediated through the efferent auditory system (Luo, Wang, Kashani, & Yan, 2008).

In the present study, in both non-musicians and musicians, the latencies of all the waves were seen to be significantly different in quiet and noise, with delay in latencies of the waves acquired in noise.The above findings are in agreement with Russo, Nicol, Musacchia and Kraus (2004) who documented the detrimental effects of noise on the subcortical representation of speech signals. The same findings were also reported by Parbery-Clark, et al., (2009).

Musicians showed significantly earlier mean latencies of all the waves than non-musicians in quiet. This is in agreement with the findings of Musacchia, et al., (2007) who found that musicians had earlier wave latencies than non-musicians in quiet. Musacchia, Stait and Kraus (2008) also documented the onset timing of musicians in quiet to be earlier than that of non-musicians. However, in contradiction Parbery-Clark, et al., (2009) found that the latencies of the waves were not significantly different in musicians and non-musicians in quiet.

In the presence of noise, the latency of wave V did not differ significantly across the groups but the latencies of the D, E and F did. In studies by Cunningham, et al., (2001), Russo et al., (2004) and Parbery-Clark, et al., (2009), it has been noted that the latency of the onset peak and transition peaks are significantly more prolonged in the presence of noise in non-musicians as compared to musicians. However, it may be noted that in the present study, the mean latencies of all the waves, including wave V are found to be earlier in musicians than in non-musicians. In agreement with this finding, Parbery-Clark, et al., (2009) had found that in noise, the onset and transition responses occurred significantly earlier in musicians than in non-musicians.

These findings indicate that long term musical training not only improves the overall encoding of temporal events of the stimuli but also restricts the detrimental effects of background noise on this process (Don & Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). The physiological basis of this finding may lie in the Medial Olivocochlear Bundle (MOCB) via which Higher-level auditory structures influence processing in lower-level structures. An increase in MOCB activity has been correlated with good speech in noise performance (De Boer & Thorton, 2008). It is possible that top-down modulation improves signal quality at the auditory periphery by selectively amplifying relevant features of the signal, and inhibiting irrelevant features in the presence of background noise. The musician's use of fine-grained acoustic information and lifelong experience with parsing simultaneously occurring melodic lines may refine the neural code in a top-down manner such that relevant acoustic features are enhanced early in the sensory system. This top-down modulation has indeed been noted to be prominent in musicians (Trainor, Shahin & Roberts, 2009) and an increase in top down modulation was been noted in children following a year musical training (Shahin, Roberts, Chau, Trainor & Miller, 2008), thus indicating the role of musical training in the sharpening of the brainstem responses in noise.

Amplitudes of Onset and Transition Peaks

The amplitudes of the onset peak V and the transition peaks D, E and F were considered for analysis. The amplitudes of the peaks are related to the robustness of the representation of the features of the stimulus (the onset and transition portions). The addition of background noise had been documented to result in reduction of amplitudes of the peaks of ABR, indicating a disruption in timing of brainstem activity (Don & Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). It has been hypothesized that the disruptive effects of noise on the representation of stimulus features may be limited by long-term musical training which can bring about enhancements of stimulus features at the subcortical level via top down influences (Dean, Harper & McAlpine, 2005) mediated through the efferent auditory system (Luo, Wang, Kashani & Yan, 2008).

In the present study, both non-musicians and musicians, the amplitudes of all the waves were seen to be significantly greater in quiet than in noise. This indicates that the noise has a detrimental effect on the subcortical representation of the signal (Don & Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). Russo, et al., (2004) and Parbery-Clark, et al., (2009) have also documented reduced amplitude of the onset and transition waves in the presence of background noise.

Musicians had higher mean amplitudes than non- musicians for all the waves in quiet, though only the amplitudes of the waves V and E are significantly different across the groups. This finding is in agreement with those of Musacchia, et al., (2007) and Parbery-Clark, et al., (2009). Parbery-Clark, et al., (2009) had documented that there were no significant differences in the amplitudes of the onset and transition waves in quiet across musicians and non-musicians, though the mean amplitudes were found to be greater for musicians.

The mean amplitudes of all the waves acquired in noise were also greater for musicians than for non-musicians, with significant differences seen in the amplitudes of waves V and D. Parbery-Clark, et al., (2009) also documented the reduction in amplitude of the onset and transition peaks in the presence of background noise to be similar in musicians and non-musicians. Though the mean amplitude of the transition wave was found to be greater in musicians, the same had not been observed with the onset wave. However, it may be pointed out that the amplitudes of onset responses are highly variable (Starr & Don, 1988; Hood, 1998) and this fact may have contributed to the differences present between the two studies.

From the above results, it is seen that the musicians have overall higher mean amplitudes of different waves in both quiet and in noise when compared to non-musicians. This is due to the disruption of the neural representation of stimulus features by noise (Russo et al., 2004). However, due to the training musicians undergo which involves the selective attention to a specific element from a complex soundscape, there is an enhanced encoding which improves the subcortical signal quality, resulting in a more robust representation of the target acoustic signal in noise. This once again points to the fact that musical training helps strengthen the sub-cortical representation of the stimulus features via top-down processes.

Formant Amplitudes

The Speech ABRs acquired from the subjects in quiet and in noise were subject to Fast Fourier Transform to obtain the amplitudes of the formants of the encoded stimulus /da/. The amplitudes of the fundamental frequency (F0), which is important for identifying the speaker, and emotional tone of voice, the first formant (F1), which provides phonetic information and the second formant (F2) were considered for analysis. It was hypothesized that that the addition of noise would result in lower formant amplitudes in the presence of noise, indicating a degradation in the neural representation of the signal.

In the present study, for both non-musicians and musicians, the amplitudes of all the formants were seen to be significantly different in quiet and noise. The mean values of the formant amplitudes were seen to be lesser in noise than in quiet. This is in line with the findings of Russo, et al., (2004) and Parbery-Clark, et al., (2009) who attributed it to the detrimental effects of noise on the neural encoding of the various formants.

The formant amplitudes were not found to be significantly different across the two groups in either quiet or in noise. This is in accordance with the findings of Parbery-Clark, et al., (2009). However, Musacchia, et al., (2007) have documented the presence of a statistically significant difference in F0 amplitude in quiet across the two groups, with musicians showing higher F0 amplitudes than their non-musically trained counterparts, though the same findings were not true of higher formants. However, it may be pointed out that in this study, musicians did show higher mean amplitudes of all formants as compared to non-musicians.

From the above findings, it was seen that both groups also showed higher mean formant amplitudes in quiet than in the presence of noise, evidence to the degradation of the neural representation of the speech signal in the presence of noise. The musicians also showed higher mean formant amplitudes than the non-musicians in both quiet and in noise, though the differences were not statistically significant. The enhanced encoding of the formants of the speech stimulus in musicians has been documented by many authors (Musacchia et al., 2007; Wong, et al., 2007). The higher mean formant amplitudes of musicians in noise as compared to non-musicians indicates a more robust sub-cortical representation of the speech signal, possibly brought about by years of continuous musical training. One possible explanation for this finding is the based on the Hebbian principle, which posits that the associations between neurons that are simultaneously active are strengthened and those that are not are subsequently weakened (Hebb, 1949). Given the present results, we can speculate that extensive musical training may lead to greater neural coherence, especially pertaining to relevant features crucial to the identification of the stimulus. This strengthening of the underlying neural circuitry would lead to a better bottom-up, feed-forward representation of the signal. We can also interpret these data within the framework of corticofugal modulation in which cortical processes shape the afferent auditory encoding via top-down processes as mentioned earlier in the discussion. Though we cannot separate the contributions of top-down and bottom-up processing, they are not mutually exclusive explanations. In all likelihood, top-down and bottom-up processes are reciprocally interactive with both contributing to the subcortical changes observed with musical training.

Comparison of Wave Latency, Wave Amplitude and Formant Amplitude Measures with Speech Perception Scores in Noise (SPIN Scores)

The performance of the subjects on the task of speech perception in noise was measured in terms of percentage correct scores on the SPIN test which used the Phonemically Balanced Wordlist in Kannada (Yathiraj & Vijayalakshmi, 2005) presented at 0dB SNR in a background of speech noise. It was speculated that the disruption of neural timing and encoding of stimulus features in the presence of competing noise would be lesser in musicians than in non-musicians, resulting in enhanced performance on the task of speech perception in noise.

The SPIN scores differed significantly across the 2 groups. Musicians had a higher mean score on the SPIN test than the non-musicians. Parbery-Clark, Skoe, Lam and Kraus (2009) and Parbery-Clark, et al., (2009) also report of a distinct advantage in musicians on the task of perception of speech in noise. This advantage was reported to correlate well with the number of years of training the musician had undergone, which strongly suggested that such intensive training helps to fine tune sensory and cognitive processes that contributed to the task of speech perception in noise.

Upon investigation as to whether the superior performance of musicians over non-musicians on the task of speech perception in noise was related to the differences in the subcortical encoding of speech stimuli across the two groups, it was found that in non-musicians, the SPIN scores did not correlate with any of the measures obtained in quiet. In musicians, the SPIN scores were found to correlate negatively with the latencies of waves D and F obtained in quiet. However, Parbery-Clark, et al., (2009) found that in quiet, there was no significant correlation between latency, amplitude or formant amplitude of brainstem responses of a subject and the corresponding scores on the task of speech perception in noise.

For waveforms acquired in noise, it was seen that for non-musicians, the SPIN scores negatively correlated with the latency of wave V and positively correlated with the amplitudes of wave V and D. Thus, poorer performance on the SPIN test was found to be related to the prolongation of onset latency and the reduction of amplitudes of the onset wave V and transition wave D in non-musicians, indicating that addition of noise had resulted in disruption of brainstem timing and a reduction in the amplitude of the responses encoding stimulus features (onset and transition), which had resulted in reduced SPIN scores. The SPIN scores of musicians correlated negatively with the latencies of wave V and F, indicating that subjects with earlier wave V and F latencies showed better performance on the task of speech perception in noise. Positive correlation was seen with the formant amplitude of F0 obtained in noise, indicating that the superior encoding of F0 in musicians had resulted in enhanced SPIN scores.

These findings indicate that musical training results in an increased resistance of the brainstem response to the disruptive effects of background noise, resulting in better timing of brainstem responses and the better encoding of stimulus features.

The findings in noise are in line with those of Parbery-Clark, Skoe and Kraus (2009) who also documented a correlation between better scores on the HINT and earlier latencies of onset and transition waves. However, the same study did not document a correlation with F0 amplitude as was seen in the present study. This may be because of the difference in the maskers used during the test of speech perception in noise. While in the present study, speech noise had been used, Parbery-Clark, Skoe and Kraus (2009) had used multi-talker babble which is a more realistic approximation of competing signals one might encounter in real life.

The higher mean SPIN scores of musicians as compared to non-musicians indicate that they have a superior ability to detect speech signals in a background of competing noise. This is a consequence of their intensive training that render them experts in extracting relevant signals from complex soundscapes. A distinct advantage is seen in musicians on the task of perception of speech in noise, which correlated strongly with the number of years of consistent practice (Parbery-Clark, et al., 2009). Musical experience was seen to result in more robust sub-cortical representation of speech in the presence of background noise, which may contribute to musician's behavioral advantage for speech in noise perception (Parbery-Clark, et al., 2009). Musicians also exhibited more faithful encoding the steady state portion of a stimulus in the presence of background noise and had higher stimulus-to-response correlations in noise than non-musicians which is indicative of more precise neural transcription of stimulus features. These enhancements may be related to the effects of the top-down (Suga, Zhang &Yan, 1997; Zhang, Suga & Yan, 1997; Luo, et al., 2008) and bottom-up processes (Hebb, 1949) that act to reduce the disruptive effects of noise while selectively enhancing stimulus features. These enhancements mean that the important features that contribute to speech intelligibility are still represented faithfully at the level of the brainstem despite the presence of a disruptive background noise. This would translate into an improved perception of speech in the presence of a competing signal.

Conclusions

Findings of this study indicates that listening and training experiences of musicians modulate their neural responses in such a manner as to allow for enhanced perception of speech stimuli in competing backgrounds.

References

Beavois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception and Psychophysics, 59*, 81-86.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language and Hearing Research*, *46*, 80-97.

Brandt, J., & Rosen, J. J. (1980). Auditory phonemic perception in dyslexia: Categorical identification and discrimination of stop consonants. *Brain and Language*, 9, 324-337.

Brattico, E., Naatanen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event related potential recordings. *Music Perception*, *19* , 199-222.

Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature, 396*, 128.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*, 1-9.

Chandrasekaran, B., Kraus, N., (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology.Â 47,* 236-246.

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: development and behavioral correlates of perception. *Ear and Hearing*, *21*, 554-568.

Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clinical Neurophysiology*, *112,* 758-767.

de Boer, J., & Thornton, A.R. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience, 28,* 4929-4937.

Don, M., & Eggermont, J.J. (1978). Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *Journal of Acoustic Society of America, 63,* 1084 -1092.

Foregeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One, 3*, 3566.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience, 17,* 1578-1592.

Galbraith, G. C., Arbagey, P.W., Branski, R, Comerci, N, Rector, P. M., (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport, 6,* 2363-2367.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neurosciences, 23* (27), 9240-9245.

Gordon-Salant, S., & Fitzgibbons, P. J. (1995). Recognition of multiply degraded speech by young and elderly listeners. *Journal of Speech and Hearing Research*, *38*, 1150-1156.

Gordon-Salant, S., & Fitzgibbons, P. J., (2004). Effects of stimulus and noise rate variability on speech perception by younger and older adults. *Journal of Acoustic Society of America, 115*, 1808-17.

Gorga, M., Abbas, P., & Worthington, D. (1985). Stimulus calibration in ABR measurements. In: The auditory brainstem response (Jacobsen J, ed), pp 49 -62. San Diego: College-Hill.

Hebb, D. O. (1949). The organization of behavior. New York: Wiley.

Hood, L. (1998). Clinical applications of the auditory brainstem response. San Diego: Singular.

Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music*, *25*, 70-83.

Klatt, D., (1980) Software for a Cascade/Parallel Formant Synthesizer. *Journal of Acoustic Society of America, 67*, 13-33.

Kraus., N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. *Annals of New York Academy of Sciences 1169*, 543-557.

Kuriki, S., Kanda, S., & Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential pianotones and chords. *Journal of Neuroscience*, *26*, 4046-4053.

Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport, 10* , 1309-1313.

Luo, F.,Wang, Q., Kashani, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience*, *28,* 11615-11621.

Micheyl,Â C.,Â Delhommeau,Â K.,Â Perrot,Â X.& Oxenham, A. J., (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219* , 36-47.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of United States of America, 104*, 15894-15898.

Musacchia, G., Strait, D., Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians. *Hearing Research, 24,* 34 -42.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., Almut., Ross., & Bernhard. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*, 169-174.

Parbery-Clark, A, Skoe, E., Lam, C., & Kraus, N (2009) Musician enhancement for speech-in-noise. Ear Hear 30:653-661.

Parbery-Clark, A, Skoe, E., & Kraus, N (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *The Journal of Neuroscience, 29* (45), 14100 -14107

Rammsayer, T., & Altenmuller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*, 37-48.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115* , 2021-2030.

Schlaug G, Jancke L, Huang Y, Staiger JF, Steinmetz H (1995a) Increased corpus callosum size in musicians.*Neuropsychologia 33*:Â 1047-1055.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H.J., Gutschalk, A.,Â &Â Rupp,A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Natural Neuroscience, 5*, 688-694.

Schon, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. Psychophysiology, *41,* 341-349.

Shahin, A., Bosnyak, D., Trainor, L., Roberts, J., & Larrey, R. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience*, *12*, 5545-5552.

Shahin, A., Roberts, L. E., & Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport, 15,* 1917-1921.

Skoe, E., & Kraus N. (2010).Hearing it again and again: on-line subcortical plasticity in humans.*PlosONE*, 5(10).

Strait, D., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research, 261*, 22-29.

Strait, D., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency: Effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29*, 661-668.

Starr, A., & Don, M. (1988). Brain potentials evoked by acoustic stimuli. In, human event-related potentials. Handbook of electroencephalography and clinical neurophysiology, Vol 3 (Picton T, ed), pp 97-158. New York: Elsevier.

Suga, N., Zhang, Y., & Yan, J. (1997). Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *Journal of Neurophysiology, 77,* 2098 -2114.

Tervaniemi, M., RytkoÂ¨nen, M., SchroÂ¨ger, E., Ilmoniemi, R. J, & NaÂ¨aÂ¨taÂ¨nen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory, 8,* 295-300

Tervaniemi, M., Kruck, S., Baene, W. D., SchrĂśger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, *30,* 1636-1642.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420-422.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonemically Balanced word list in Kannada*. Developed in Department of Audiology, AIISH, Mysore.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science, 12*, 732-745.