AUDITORY BRAINSTEM RESPONSES TO FORWARD AND REVERSED SPEECH IN NORMAL HEARING INDIVIDUALS

*Sujeet Kumar Sinha, **Vijayalakshmi Basavaraj

Abstract

Differences in the coding of forward and reversed speech has indicated that the human auditory system is sensitive to different types of speech sounds. Infants as well as adults are reported to respond differently to forward and reversed speech. Functional magnetic resonance imaging (FMRI) have revealed that listening to forward speech activates large regions of the temporal lobe, whereas reverse speech evokes significantly diminished and nonlocalised brain responses. The objective of the present study was to assess the differences, if any, in the brainstem responses to forward and reversed speech stimuli. 50 normal hearing adults participated for the study. A synthesized 40msec short stimulus /da/ syllable was used as the stimulus for both forward and reversed conditions. The syllable was reversed with the help of Adobe Audition software. Auditory brainstem responses were recorded for the forward and reversed /da/ stimulus. Results revealed that the amplitude of wave V was larger for reversed speech as compared to the forward speech. On the other hand, the amplitude of the frequency following responses, fundamental frequency and the formant frequency were smaller in the reversed speech condition as compared to the forward speech condition. The findings of the present study suggest that differential processing of forward and reversed speech occurs at the brainstem level as well even for a short duration stimulus. The better response to forward speech could be due to the universal temporal and phonological properties of human speech which is familiar to the brainstem and hence is processed efficiently. These findings suggests that Speech evoked ABR may throw light to understand the encoding of complex acoustic stimulus at the brainstem level.

Key words: Brainstem, Speech Evoked ABR, Forward Speech, Reversed Speech,

The auditory cortex plays a major role in the perception of speech, music and other meaningful auditory signals. Subcortical processing of sounds dynamically interacts with cortical processing to reflect important nonsensory factors such as musical expertise (Musacchia, Sams, Skoe & Kraus, 2007; Wong, Skoe, Russo, Dees & Kraus, 2007), linguistic experience (Krishnan, Xu, Gandour & Cariani, 2005), and attention (Galbraith, Bhuta, Choate, Kitahara & Mullen, 1998; Galbraith, Olfman & Huffman, 2003).

One of the ways to investigate the nature of auditory brainstem processing is to record speech evoked auditory brainstem responses (sABR), since sABR consist of transient and a sustained portion. The sustained portion is also known as frequency following responses (FFR). It provides a direct electrophysiological measure of sensory processing at the subcortical levels of the auditory pathway (Galbraith et al. 2000). It has been reported that short-term and long-term auditory experiences initiated in childhood could alter brainstem processing. For example, Russo, Nicol, Zecker, Hayes, and Kraus (2005) found improved auditory brainstem timing to speech stimuli in background noise in children with language-based learning problems following an 8-week auditory speech training program. Language and musical experience to influence auditory encoding of sound at subcortical levels of the auditory pathway suggesting that these areas are more plastic and dynamic than was typically assumed to be by sensory neuroscientists, and that at least some of these influences are opined to be mediated by the

*Lecturer, Department of Audiology, All India Institute of Speech & Hearing (AIISH), Mysore, e-mail: sujitks5@yahoo.co.in,
**Director, AIISH, Mysore, e-mail: vbasavaraj@hotmail.com

top-down mechanisms (Banai & Kraus, 2008).

Given the information that the brainstem does contribute to processing of speech, it will be interesting to study the nature of processing of the different aspects of speech by the subcortical structures.

Reversed speech is considered as a good stimulus to investigate in this area since reversed speech has the unique characteristics to maintain the physical characteristics of speech such as the distribution of frequency of sounds, their global amplitude and, to some extent, their temporal and rhythmic characteristics. The main difference between forward speech and reversed speech lies in the coarticulations which are totally distorted in the reversed signal. If speech stimuli are played backwards it sounds like unfamiliar and often bizarre sounding language eventhough phoneme duration and the fundamental voicing frequency are preserved (Binder et al.2000; Dehane, Dehane & Hertz-Pannier, 2002). This is because reverse stimulation violates phonological properties universally observed in human speech (Binder et al., 2000; Dehane et al., 2002).

There are reports which suggest differences in the coding and processing of forward and reverse speech stimulus. Adults as well as infants are sensitive to these stimulus differences. It has been reported that 4 days old neonates and 2 months old infants can discriminate native and foreign languages but not when those sounds are played backwards (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988). Functional magnetic resonance imaging (FMRI) have indicated that the left angular gyrus, right dorsolateral prefrontal cortex and the left mesial parietal lobe (precuneus) get significantly more activated by forward speech than by backward speech. (Mehler et al., 1988). FMRI studies have also shown that listening to forward speech activates large regions of the temporal lobe, but backward speech evokes significantly diminished and nonlocalised brain responses (Binder et al. 2000). The direct contrasts Words-Pseudowords and Words-Reversed have no areas of significant activation difference in either direction in neither hemisphere, nor a direct contrast between Pseudowords and Reversed conditions (Binder et al., 2000).

The segmental and suprasegmental features of speech may condition and modify brainstem neurons to process familiar sounds more selectively and preferentially. It is also possible that the type of signal processing may affect the subsequent cortical development and language lateralization. Galbraith et al. (2004), where they have obtained the brainstem responses to forward and reversed speech (using an 880 msec 4 syllable phrase. Both horizontal and vertical electrode montages were used to record the responses on a small sample of 11 subjects. There is a need to ascertain the findings of Galbraith et al. (2004) on a larger sample. Also, there is a need to explore the auditory brainstem processing using different types and duration of speech stimuli, with and without background noise and with different methods of presentation. The information documented by such studies especially using a larger sample than that used by Galbraith et al (2004) would throw light on the similarities and differences between the subcortical and cortical processing of speech, the interaction between the two levels and implication of these interactions or the lack of it. Since, Galbraith et al (2004) did only FFT analysis of the brainstem responses, it will be interesting to measure the amplitude of each peak to substantiate the differences between the forward and reversed speech. Therefore there is a need to study the amplitude of each peak and FFT analysis on a larger sample. Thus, the objective of the present study was to assess the possible differences in brainstem responses to forward and reversed speech stimuli.

Method

Research design: "A Within Subject" research design was used where in the responses of each subject to forward and reversed speech stimuli were compared.

Hypothesis: The null hypothesis that there is no difference between the ABR responses for forward and reversed speech in subjects with normal hearing sensitivity was adopted.

Participants:

Fifty young adult students (30 males and 20 females) in the age range 17 to 23 years, with a mean age of 19 years consented to participate in the study. All the subjects had normal hearing thresholds as defined by puretone thresholds of <20 dBHL from 250 Hz to 8000 Hz with normal middle ear functions as revealed by A type of tympanograms and presence of acoustic reflexes present at 500 Hz, 1000 Hz, 2000 Hz & 4000 Hz for both ipsi and contralateral stimulation.

Test Stimulus:

Synthesized /da/ syllable of 40 msec length was used as the test stimulus, synthesized with a Klatt synthesizer (Klatt, 1980). The stimulus was prepared to include an onset burst frication at F3, F4, and F5 during the first 10 msec and a fundamental frequency range of 105-121 Hz, followed by 30-msec F1 and F2 transitions ceasing immediately before the steady-state portion of the vowel. Although the stimulus does not contain a steady-state portion, it is psychophysically perceived as a consonant-vowel speech syllable. Such a stimulus was first developed at Northwestern University by King et al (2002) and the same has been used for research at Northwestern University.

Figure- 1 shows the /da/ stimulus of 40 msec whereas the figure 2 shows the reversed waveform of the same stimulus. Stimulus in the figure-2 is the mirror image of the stimulus in figure-1. Adobe audition version-2 software was used to reverse the stimulus.

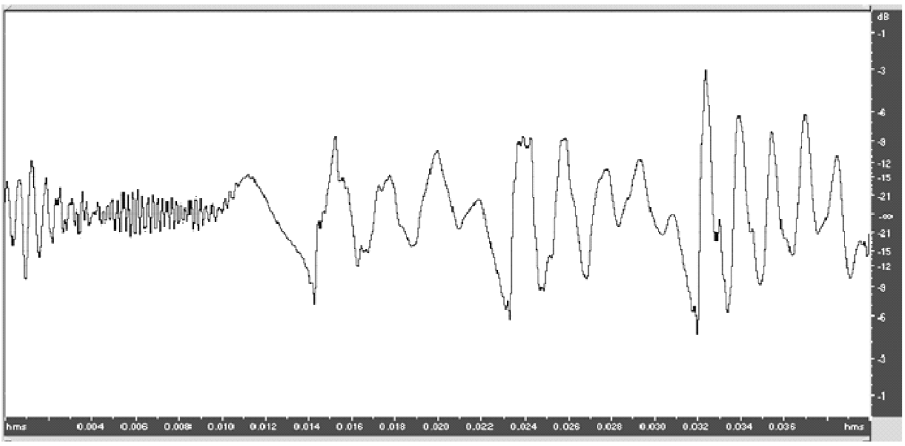

Figure. 1. Waveform of the forward /da/ stimulus.

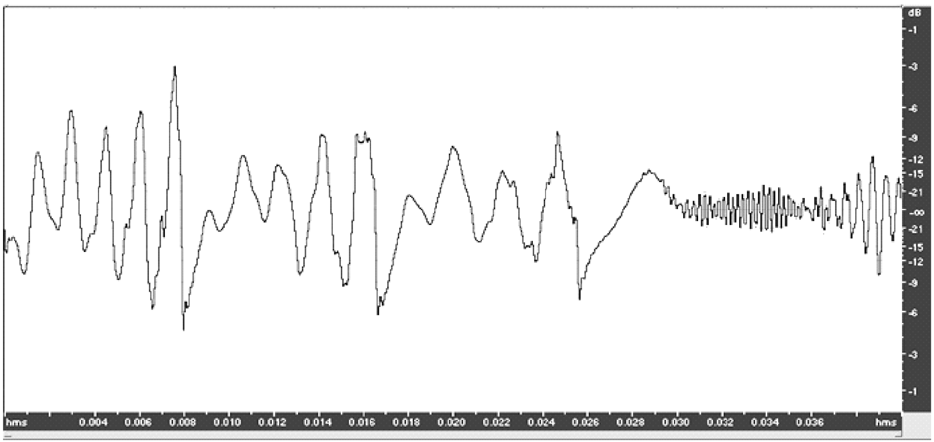

Figure.2. Waveform of the temporally reversed /da/ stimulus.

Instrumentation:

- A calibrated (ANSI S3.6-1996), two channel clinical audiometer Madsen OB922 with TDH-39 headphones housed in Mx-41/AR ear cushions were used for Pure tone audiometry. Radioear B-71 bone vibrator was used for measuring bone conduction thresholds.
- A calibrated middle ear analyzer, (GSI Tympstar) using 226 Hz probe tone was used for tympanometry and reflexometry.
- Intelligent Hearing (Smart EP windows USB version 3.91) evoked potential system with insert ear ER-3A receiver was used for recording auditory brainstem responses.

Procedure

All the subjects underwent puretone audiometry and tympanometry to ensure that they had normal hearing sensitivity and normal middle ear functions. The speech evoked auditory brainstem responses were recorded for all the subjects for both the forward and the reversed /da/ stimulus in the EP system of Intelligent Hearing systems version 3.91. The details of the protocol for recording the speech evoked ABR are given in table-1

Analysis:

Speech evoked ABR consists of six peaks. These peaks are labeled as (V, C, D, E, F, & O. (Russo et al., 2004; Russo et al., 2005, Johnson et al., 2005), The amplitude of waves V, D, E and F were measured for the forward as well as reversed conditions. Wave C and wave O were not taken into consideration for analysis as they were not present in all the subjects. Two audiologists who have the knowledge of the Speech evoked ABR analyzed the waveforms independently. The inter audiologist reliability was ensured by doing a correlational analysis for all the peaks. All the peaks showed a high positive correlation for the peaks marked by the two audiologists.

Measurement of Fundamental frequency (F0) and First Formant frequency (F1):

FFR consists of energy at fundamental frequency of the stimulus and its harmonics. The period between response peaks D, E, and F in the recorded waveform corresponds to the wavelength of the F0 of the utterance (Johnson et al., 2005). Moreover, Fourier analysis of this portion of the response confirms a spectral peak at the frequency of F0. Additionally, the spacing of the small, higher-frequency fluctuations between waves D, E, and F correspond in frequency to the F1 of the stimulus (Russo et al., 2004; Russo et al., 2005, Johnson et al., 2005). Fast Fourier analysis was performed on the recorded waveform. Activity occurring in the frequency range corresponding to the fundamental frequency (F0) of the speech stimulus (103-121Hz) and activity corresponding to the first formant frequency (F1) of the stimulus (220 Hz -729 Hz) were measured. 2 ms on-2 ms off Hanning ramp was applied to the waveform. Zero-padding was employed to increase the number of frequency points where spectral estimates were obtained. A subject's response was required to be above the noise floor in order to include in the analysis. This calculation was performed by comparing the spectral magnitude of pre stimulus period to that of the response and if the quotient of the magnitude of the F0 or F1 frequency component was greater than or equals to one the response was considered to be present. The analysis of F0 and F1 was done with the MATLAB software.

Results

A long term average speech spectrum of both the forward and reversed speech was performed to see whether the spectrums of the two sounds are different. On analysis it was found that the spectrum of the forward and reversed stimuli remained the same. Figure 3 shows the long term average spectrum of the forward and reversed speech stimuli. Since there was a perfect overlap of the two spectra it was difficult to differentiate one from the other. Hence, the SPL of the reverse speech (shown in continuous line) was deliberately reduced to differentiate it from the forward speech spectrum (shown in dotted line).

Wave V was identified for the forward and the reversed speech similar to the way it is identified for click stimulus. Since wave V is the result of an onset response, this is similar to both the click and the speech evoked ABR. Johnson et al., (2005) have reported and illustrated that the visual analysis of /da/ stimulus waveform and its corresponding brainstem response has several similarities. They

Stimulus parameter	Stimulus	Forward /da/ and reversed /da/
	Duration	40 msec
	Intensity	80 dB nHL
	Polarity	Alternating
	Repitition rate	9.1/sec
	Total no. of stimulus	2000
Acquisition parameter	Analysis time	60 msec
	Filter setting	30 to 3000 Hz
	Electrode Montage	Noninverting(+ve):Vertex, Inverting(--ve): Test ear mastoid, Ground: Non test ear mastoid
	Transducer	Insert receiver-ER-3A

Table 1: Recording parameters for the speech evoked auditory brainstem responses for the forward /da/ and the reversed /da/ stimulus

have recommended to shift the stimulus waveform by approximately 7 msec to account for neural conduction time to identify the speech ABR peaks which correlate with the peaks in the stimulus, namely peak D, E and F. In the present study the speech ABR peaks corresponding to the peaks D, E and F in the stimulus were identified using the same procedure for both the forward and reversed speech keeping the burst of the stimulus as reference. The burst for the forward stimulus appears in the beginning of the stimulus and for the reversed speech it appears in the last of the stimulus and hence, the peaks D, E, and F occurs in the reversed order for the reversed speech and thus it can be seen from figure 4b that wave F follows wave V against wave D.

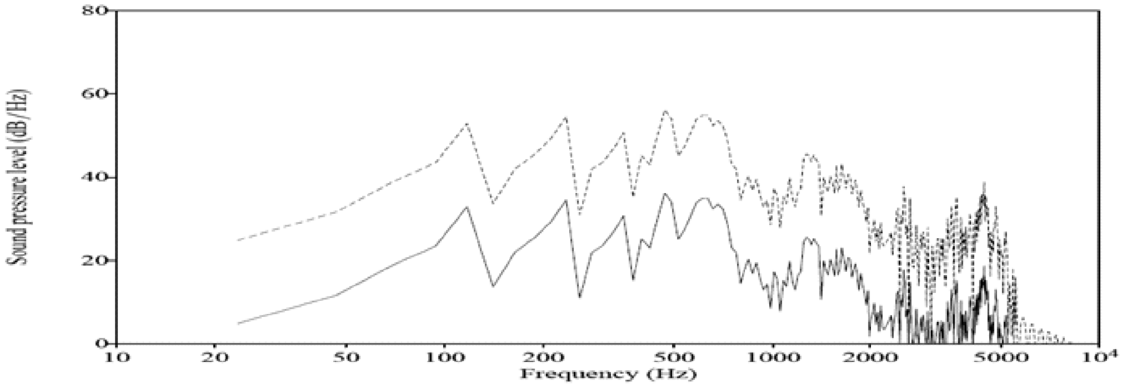

Figure- 3. Long term average speech spectrum of forward and reversed speech

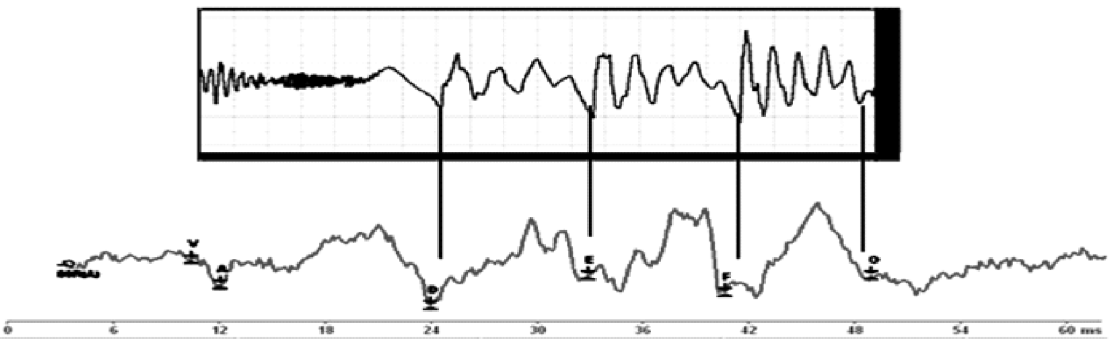

Figure 4a. Sample of Speech evoked ABR for the forward Speech and its correlation with the stimulus peaks

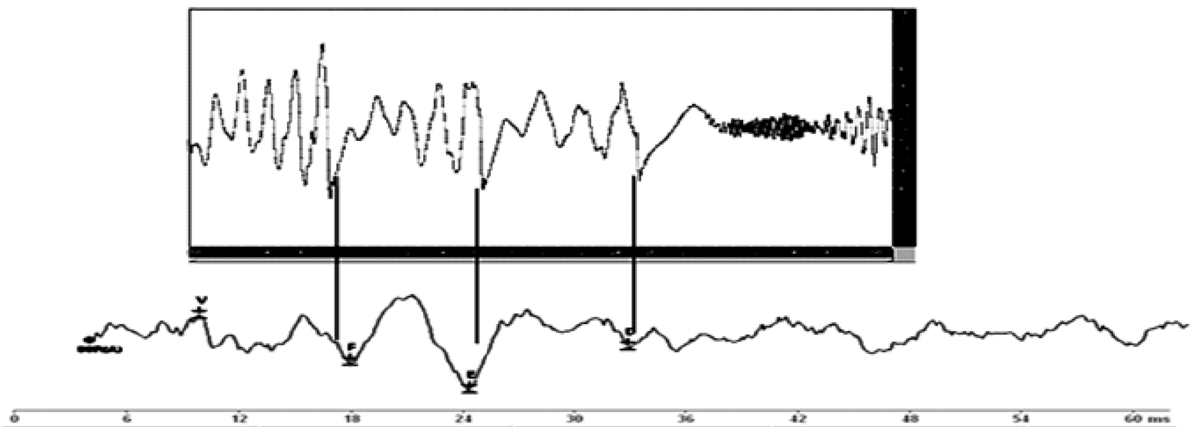

Figure 4b. Sample of Speech evoked ABR for the reversed Speech and its correlation with the stimulus peaks

Descriptive statistics:

SPSS software version 15 was used for statistical analysis. Mean and standard deviations for the amplitude alone of waves V, D, E and wave F were determined for all the subjects for the forward and reverse speech. Latency parameter was not subject to analysis as this is determined by the stimulus parameters. The mean and Standard deviation (SD) of amplitude of the different waves for the forward & reversed speech are presented in table 2

From the table 2 it can be noticed from that the mean amplitude for wave V is larger for the reversed speech as compared to that for the forward speech condition. The amplitude of others peak (waves D, E & F) are larger for forward speech condition as compared to that for the reversed speech condition. This can be seen in figure 5 as well.

Wave	Forward Speech Amplitude (µv)		Reversed Speech Amplitude (µv)	
	Mean	SD	Mean	SD
Wave V	0.18	0.10	0.33	0.18
Wave D	0.33	0.25	0.22	0.14
Wave E	0.35	0.16	0.22	0.14
Wave F	0.28	0.12	0.20	0.09

Table 2: Mean and standard deviations (SD) for amplitude (µv) of different peaks for the forward & reversed speech

To know the significance of difference between the amplitude of different peaks the dependent 't' test was done. The results of dependent 't' test revealed a significant difference between the amplitude of wave V [t (49) = 6.54, p

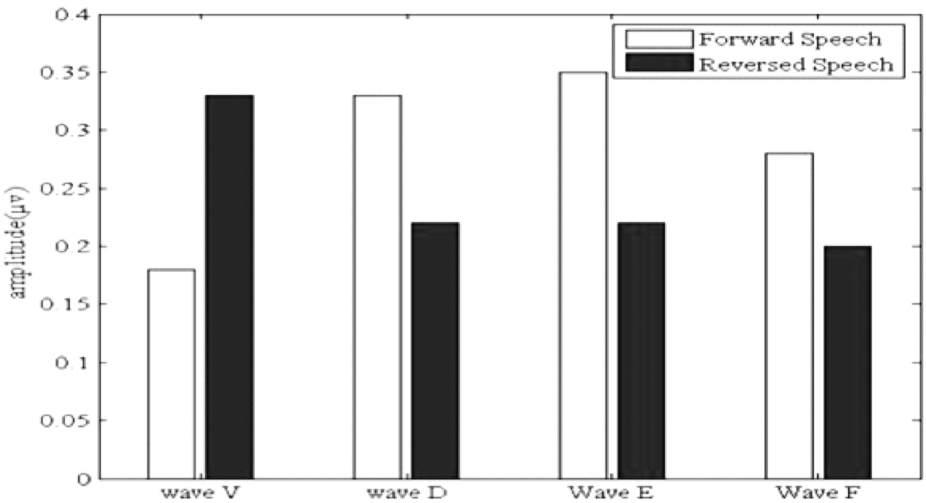

Figure 5. Amplitude (µv) of different peaks of forward & reversed Speech

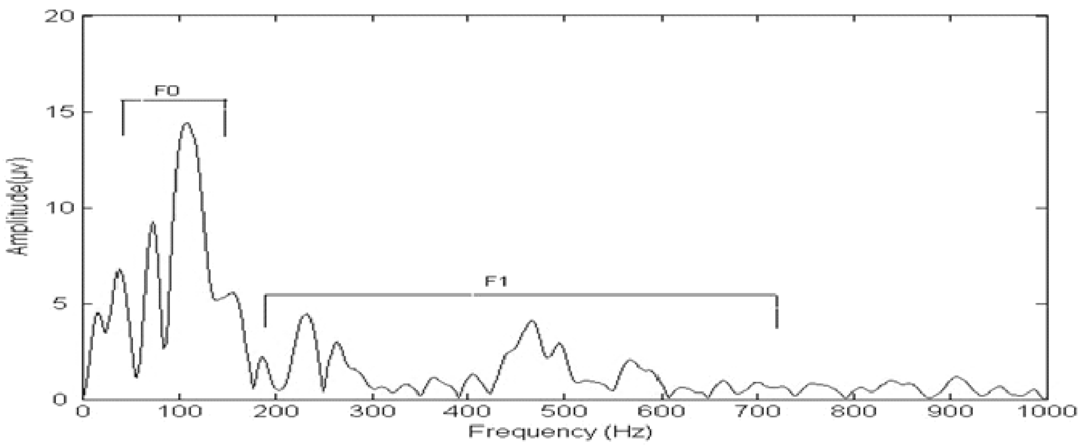

Figure 6. Analysis of F0 and F1. Response indicates that only the fundamental frequency and first formant frequency (F0=103-121 Hz; F1= 220 Hz to 720 Hz) were measurable (For Forward Speech).

< 0.01], wave D [t (49) = 6.27, p < 0.01], wave E [t (49) = 5.03, p<0.01] & wave F [t (49) = 4.80, p<0.01]. Also an analysis of fundamental frequency and first formant frequency was done. This was done using "Matlab Software". A sample figure of how the fundamental frequency and the first formant frequency were measured is given in the figure 6 above.

F0 and F1 analysis:

Analysis of Fundamental frequency and first formant frequency revealed that the mean amplitude of fundamental frequency for forward speech was 15.72µv and of the reversed speech was 9.23 µv. The amplitude of first formant frequency for forward speech was 12.86 µv and that of reversed speech was 7.83 µv. A dependent't' test was applied to compare the amplitude of fundamental frequency and the first formant frequency of forward speech and reversed speech. This revealed a significant difference for the F0 [t (49) = 2.34, p<0.05] and for the F1 [t (49) =2.22, p<0.05].

Thus the null hypothesis that there is no significant difference between the brainstem responses to forward and reversed speech conditions was rejected. To summarize the results, amplitude of all the peaks for forward speech except for "wave V" was more as compared to that of the reversed speech. Also, the amplitude of fundamental frequency and first formant frequency was more for forward speech as compared to the reversed speech.

Discussion

Auditory brainstem is the site of extensive synaptic complexity and acoustic signal processing in the auditory pathway (Eggermont, 2001). The regularities in the acoustic biotope, consisting of individual vocalizations and background sounds that are part of the natural habitat are likely to be manifested in the response properties of auditory neurons (Aertsen, Smolders, & Johannesma, 1979; Nelken, Rotman, & Yosef, 1999; Smolders, Aertsen, Johannesma, 1979).

In the present study the frequency following responses to forward and reversed speech were recorded in a vertical montage. The vertical frequency following responses measure responses originating in the rostral brainstem (Galbraith, 1994; Gardi, Merzenich & Mckean, 1997). The prinicipal finding of the present study is that there is a significant difference between the forward and the reversed speech coding even for a short duration stimulus of 40 msec at the brainstem level. The amplitude of the frequency following responses to reversed speech stimulus were reduced as compared to that for the forward speech. Further, the results also indicate that the amplitude of fundamental frequency and the first formant frequency were also reduced in the individuals for the reversed speech condition. The reduced amplitude of FFR, fundamental frequency and the first formant frequency suggests that the brainstem processes the forward speech differently than the reversed speech. Galbraith et al (2004) have also reported a reduced FFT response to reversed speech compared to the forward speech. The present study supports their findings and further illustrates that the differential processing is seen even for a short duration CV stimulus like /da/. However, responses obtained for short duration stimuli using horizontal montage needs to be explored as present study used only vertical montage.

It is possible that the reduced amplitude of the frequency following responses (i.e. the amplitude of the wave D, E and F) may simply be due to the coarticulation effect in forward and reversed speech. The coarticulations are reported to be totally distorted in the reversed signal (Binder et al., 2000; Dehane, et al., 2002). One may argue that the reduced responses in the brainstem may be due to the distortion of the coarticulations in the reversed speech rather to the differences in the processing at brainstem level. However in the present study, the FFT analysis of the FFR shows reduced amplitude of F0 and F1. It is difficult to explain the reduced responses of F0 to the coarticulation effect because in the reversed speech some of the parameters such as distribution of frequency of sounds, their global amplitude, phoneme duration and the fundamental voicing frequency are preserved (Binder et al.2000; Dehane, et al. 2002), as shown in figure 3 also. Therefore, the findings in the present study may not be due to the distortion of the coarticulation effect alone. It appears to be because of the differential processing of forward and reversed speech at the brainstem level. Thus, the results of the present study suggest that the brainstem structures processing is also different for familiar and non familiar stimuli similar to the cortical processing.

It appears that the synaptic processing at the level of rostral brainstem is more effective for speech stimuli characterized by highly familiar prosodic and phonemic structures as illustrated by better ABR responses for the forward speech condition. This could be due to the conditioned exposure to native speech patterns that may modify the microanatomy and processing capabilities of the auditory system (Querleu & Renard, 1981). Indeed there are studies which suggest that there is plasticity even at the level of brainstem (Russo et al, 2005). The notion that neural activity in the rostral brainstem is sensitive to language experience, (i.e., language-dependent) is also reported (Krishnan et al. 2005). At this point, a question arises as to whether these observed FFR effects are stable for all types of stimuli. Further studies with a longer duration stimuli, a tonal stimulus, words and sentences will strengthen the present area of research.

Conclusion

The present study highlights the differential processing of forward and the reversed speech at the brainstem level similar to that at the cortex. Differences in the processing at the cortical level for forward and reversed speech has been reported (Binder et al., 2000). Findings of the present study suggest that the differential processing occurs at the brainstem level as well. The differences in the processing of forward (familiar) and reversed (non familiar) speech could be due to the previous exposure to the forward speech making the universal temporal and phonological properties of speech familiar to the auditory system. Findings of the present study also suggest that speech evoked ABR provides information to understand the encoding of complex acoustic stimulus at the brainstem level. Further research on normal and abnormal speech evoked ABR may throw light on some of the factors contributing to the poor speech perception in the clinical population. Although speech perception involves various cognitive processes that go beyond a single neural code and the brainstem encoding of speech sounds alone may not account for the speech perception, it is possible that abnormal neural response patterns at the brainstem may be one of the many factors which contributes to the poor speech perception.

References

Aertsen, A.M., Smolders, J.W., & Johannesma, P.I., (1979). Neural representation of the acoustic biotope: on the existence of stimulus event relations for sensory neurons. *Biological cybernetics, 32,* 175-185

Banai, K., & Kraus, N. (2008). The dynamic brainstem: implications for APD. In: D. McFarland and A. Cacace (eds). *Current Controversies in Central Auditory Processing Disorder* (pp.269-289). Plural Publishing Inc: San Diego, CA.

Binder, J., Frost, J., Hammeke, T., Bellgowan, P., Springer, J., Kaufman, J., et al. (2000). Human temporal lobe activation by speech and non speech sounds. *Cerebral Cortex, 10, 512-528.*

Dehaene, S., Deahene, G., & Hertz-Pannier, L. (2002). Functional neuroimaging of Speech perception in infants. *Science, 298,* 2013-2015.

Eggermont, J. (2001). Between sound and perception: reviewing the search for a neural code. *Hearing Research, 157,* 1-42.

Galbraith, G. C., Bhuta, S. M., Choate, A. K., Kitahara, J. M., & Mullen, T. A., Jr. (1998). Brain stem frequency-following response to dichotic vowels during attention. NeuroReport, 9, 1889-1893.

Galbraith, G. C., Olfman, D. M., & Huffman, T. M. (2003). Selective attention affects human brain stem frequency-following response. *NeuroReport, 14,* 735-738.

Galbraith, G. C., Threadgill, M. R., Hemsely, J., Salour, K., Songdej, N., Ton, J., et al. (2000). Putative measures of peripheral and brainstem frequency following frequency following in humans. *Neuroscience Letters,* 292, 123-127.

Galbraith, G. C., (1994). Two channel brainstem frequency following responses to puretone and missing fundamentals stimuli. *Electroencephalography and Clinical Neurophysiology,* 92, 321-330.

Gardi, J., Merzenich, M., & McKean, C. (1979). Origins of the scalp recorded frequency following responses in the cat. *Audiology, 18,* 353-381.

Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brain stem response to speech: A biological marker of auditory processing. *Ear and Hearing, 26,* 424-434

King, C., Catherine M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neuroscience Letters, 319, 111-115.

Klatt, D. (1980). Software for cascade / parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971-5.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research and Cognitive Brain Research, 25,* 161-168.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. A.(1988). A precursor of language acquisition in young infants. *Cognition, 29,* 143-178.

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 Science USA, 104,* 15894-15898.

Nelken, I., Rotman, Y., & Yosef, O.B. (1999). Responses of auditory cortex neurons to structural features of natural sounds. *Nature, 397,* 154-157.

Querleu, D., & Renard, X. (1981). Les perceptions auditives ju foetus humain. *Medical Hyg, 39,* 2012-2110.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115,* 2021-2030

Russo, N., 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.

Smolders, J.W., Aertsen, A.M., & Johannesma, P.I., (1979). Neural representation of the acoustic biotope. A comparison of the response of auditory neurons to tonal and natural stimuli in the cat. *Biological cybernetics, 35,* 11-20.

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.

Acknowledgements

- Director, All India Institute of Speech and Hearing, Mysore
- Head of the Department of Audiology for permitting to conduct the study.
- Dr. Nina Kraus and Dr. Erika Skoe of Northwestern University for permitting us to use the /da/ stimulus for the study.
- All the students who participated in the study