

Speech Evoked Auditory Brainstem Responses: A New Tool to Study Brainstem Encoding of Speech Sounds

Sujeet Kumar Sinha · Vijayalakshmi Basavaraj

Received: 19 February 2009 / Accepted: 28 April 2010 / Published online: 11 January 2011
© Association of Otolaryngologists of India 2011

Abstract The neural encoding of speech sound begins in the auditory nerve and travels to the auditory brainstem. Non speech stimuli such as click or tone bursts stimulus are used to check the auditory neural integrity routinely. Recently Speech evoked Auditory Brainstem measures (ABR) are being used as a tool to study the brainstem processing of Speech sounds. The aim of the study was to study the Speech evoked ABR to a consonant vowel (CV) stimulus. 30 subjects with normal hearing participated for the study. Speech evoked ABR were measured to a CV stimulus in all the participants. The speech stimulus used was a 40 ms synthesized/da/sound. The consonant and vowel portion was analysed separately. Speech evoked ABR was present in all the normal hearing subjects. The consonant portion of the stimulus elicited peak V in response waveform. Response to the vowel portion elicited a frequency following response (FFR). The FFR further showed a coding of the fundamental frequency (F0) and the first formant frequency (F1). The results of the present study throw light on the processing of speech in brainstem. The understanding of speech evoked ABR has other applications both in research as well as in clinical purposes. Such understanding is specially important if one is interested in studying the central auditory system function.

Keywords Auditory brainstem responses · Speech evoked ABR · Transient response · Frequency following responses

Introduction

The acoustic signal is transmitted through external and middle ear mechanically to the cochlea is transduced into electrical impulse for further transmission via the auditory nerve. This is reported to be highly complex transformation. It is well established that the primary role of the cochlea is separation of spectral information prior to further transmission via the auditory nerve. The central auditory system (CAS) is composed of a complicated network of projections between the cochlea, multiple brainstem nuclei, the mid-brain, the thalamus, and the cortex.

An essential function of the central auditory system is the neural encoding of speech sounds. The neural encoding of the sound begins in the auditory nerve and travels to the auditory brainstem [1]. In the clinical practice brainstem responses to simple stimuli such as click or tones are widely used to assess the integrity of the auditory pathway [2]. However a non-speech stimulus such as click or tone does not provide insight about the actual processing of speech sounds. The study of encoding of speech sound at the brainstem level is essential since speech is a complex signal varying in many acoustic parameters over time. A complete description of how the auditory system responds to speech can only be obtained by using speech stimuli.

Recently Speech evoked auditory brainstem responses measures have been introduced as a means to study the brainstem encoding of speech sounds [1, 3]. It has been established as a valid and reliable means to assess the integrity of the neural transmission of speech stimuli at the

S. K. Sinha (✉)
Department of Audiology, All India Institute of Speech and Hearing, Mysore 570006, India
e-mail: sujitks5@yahoo.co.in

V. Basavaraj
All India institute of Speech and Hearing, Mysore, India

brainstem. Typically with a click or a tone stimulus we get an auditory brainstem response which consists of mainly peaks I, III, and V, whereas with a consonant vowel syllable (CV) the auditory brainstem responses (ABR) consist of responses which to consonant and vowel separately.

To a consonant–vowel syllable, the onset response (peaks V, A) represents the burst onset of a voiced consonant, whereas later portions are likely to represent the offset of the onset burst or the onset of voicing (wave C) and the offset of the stimulus (wave O). The harmonic portion of the speech stimulus gives rise to the frequency-following response (FFR, waves D, E, and F) [1, 4–6]. Also the period difference between the response peaks D, E, F corresponds to the wavelength of the Fundamental frequency (F0) and a Fourier analysis of this portion of the response confirms a spectral peak at the frequency of F0 and also a spectral peak at first formant frequency (F1) [1, 6].

In recent studies speech evoked ABR has brought an insight into the diagnosis of children with learning disability [3, 5–7]; an insight into studying the developmental plasticity of human brainstem to speech sounds [8]; understanding tonal language processing skills [9, 10], studying temporal encoding of amplitude modulations [11], and has also shown to be a means for evaluating training related improvements in children with learning disability [12]. Studying the processing of speech sounds at the brainstem may provide knowledge regarding the central auditory processes involved in normal hearing individuals and also in clinical populations. The knowledge of brainstem processing of speech sounds may be applied in understanding effects of peripheral hearing loss on processing of speech sounds, in individuals with learning problems, children with auditory processing problems, hearing aids, cochlear implants, plasticity of the auditory system, presbycusis or auditory neuropathy. In the Indian context, in a reported study [13] Speech evoked ABR was elicited using a dichotic paradigm. The authors reported that dichotic stimulus resulted in two different response patterns between the right and left channel. There is a dearth of studies in India and hence the present study was taken up to document Speech evoked ABR among the Indian population and to further understand the brainstem processing of Speech sound.

Methods

Subjects: Thirty subjects in the age range of 18–25 years participated for the study. These participants were the students of All India Institute of Speech & Hearing, Mysore. Subjects were selected on the basis of normal hearing sensitivity and a normal middle ear function on the basis of Puretone audiometry and immittance audiometry

findings. All the participants had hearing threshold at or below 15dBHL for octaves from 250 to 8000 Hz and ‘A’ type tympanogram with the ipsilateral and contralateral reflexes present in both the ears. All the participants were informed about the purpose of the study and consent was taken before the study.

Instrumentation

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

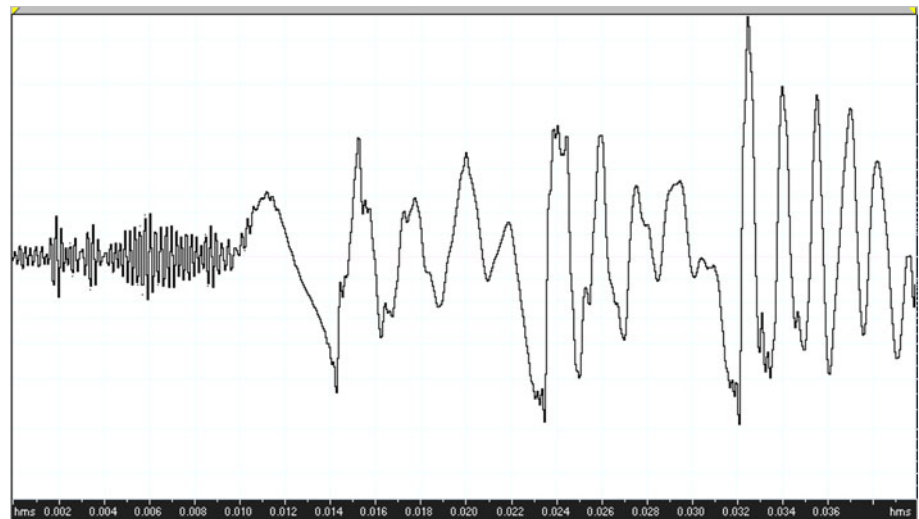
Stimulus and Recording Parameters

Stimulus to record the speech evoked ABR was the stop consonant/da/. This particular stimulus has been used by Dr. Nina Kraus at Northwestern University to study Speech evoked ABR. A Klatt cascade/parallel formant synthesizer has been used to synthesize a 40-ms speech-like/da/syllable at a sampling rate of 10 kHz (Fig. 1). The stimulus was constructed to include an onset burst frication at F_3 , F_4 , and F_5 during the first 10 ms, followed by 30-ms F_1 and F_2 transitions ceasing immediately before the steady-state portion of the vowel. The stimulus did not contain a steady-state portion.

The stimulus was presented monaurally with alternating polarity at 80 dB nHL via insert ear phone, with a repetition rate of 7.1/s Responses were collected with silver chloride electrode, differentially recorded from Cz (active) to ipsilateral earlobe (reference), with the opposite ear mastoid as ground. A total of 2000 sweeps were used for recording speech evoked ABR. Waveforms were averaged in Intelligent Hearing Systems (IHS) with a recording time window spanning 10 ms prior to the onset and 20 ms after the offset of the stimulus. A filter setting of 30–3000 Hz was used for recording.

Measurement of Fundamental Frequency (F0) and First Formant Frequency (F1)

FFR consists of energy at fundamental frequency of the stimulus and its harmonics [14]. Fast Fourier analysis was performed on the 11.5–46.5 ms of the recorded waveform.

Fig. 1 Stimulus waveform of/da/stimulus

Activity occurring in the frequency range corresponding to the fundamental frequency (F0) of the speech stimulus (103–121 Hz) and activity corresponding to the first formant frequency (F1) of the stimulus (220–729 Hz) was also measured. A subject's response was required to be above the noise floor in order to include in the analysis. This calculation is 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 is considered to be present.

Results

Based on evaluations of 30 subjects the speech evoked ABR latency and amplitudes of different parameters were analysed. Speech evoked ABR consisted of two parts: (I) the transient portion and (II) sustained portion which is also

known as frequency following response (FFR). The response to the onset of the speech stimulus included a positive peak (wavelet V), followed by a negative trough (wavelet A). Following the onset response peaks 'C' 'D' 'E' 'F' present in the FFR portion. There was a peak to the offset of the stimulus which is named as 'O'. The present study analysed the major peaks of speech evoked ABR. Sample waveform of one of the subjects has been given in Fig. 2.

The mean and SD of the latency and amplitude of different peaks of Speech evoked ABR has been given in the Table 1.

Measurement of Amplitude of Fundamental Frequency and First Formant Frequency

Measurement of the amplitude of F0 and F1 indicated that the mean amplitude of F0 was more than the mean amplitude of F1. The values of the amplitude of the fundamental and first formant frequency are given in Table 2

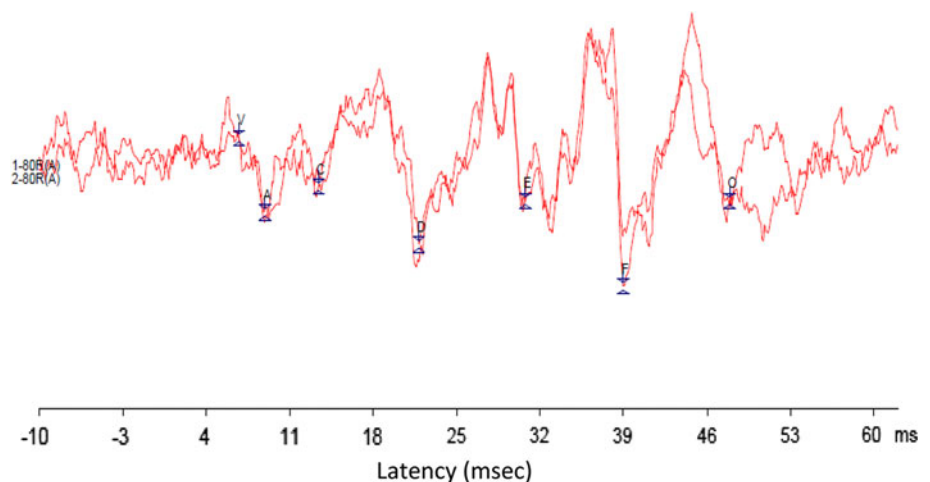
Fig. 2 Speech Evoked ABR recorded in one of the subjects

Table 1 Mean & SD of the Latency and amplitude of different peaks of speech evoked ABR

Peaks	Latency (ms)		Amplitude (μV)	
	Mean	SD	Mean	SD
V	6.81	0.44	0.19	0.11
C	16.82	1.93	0.24	0.16
D	24.75	1.02	0.32	0.23
E	31.36	0.77	0.37	0.17
F	40.04	1.09	0.29	0.19

Table 2 Mean and Standard deviation (SD) of amplitude of F0 and F1

Parameters	Mean (μV)	SD (μV)
Fundamental frequency	17.16	7.88
First formant frequency	8.83	3.06

The sample waveform of the analysis of fundamental frequency and first formant frequency are shown in Fig. 3

Discussion

Speech evoked ABR provides reliable information about the speech coding at the brainstem level. Speech evoked ABR may be used to characterize the specific measures of brainstem function that may be used to characterize neural encoding of speech sounds for clinical and research applications. The two measures in Speech evoked ABR that is the transient and frequency following responses provides information regarding auditory pathway encoding of consonant and vowel portion of the stimuli. The study described explicit method to record a speech evoked ABR to understand the neural basis of speech encoding.

The transient portion of the speech evoked ABR for/da/ sound showed an initial wave V present at a mean latency

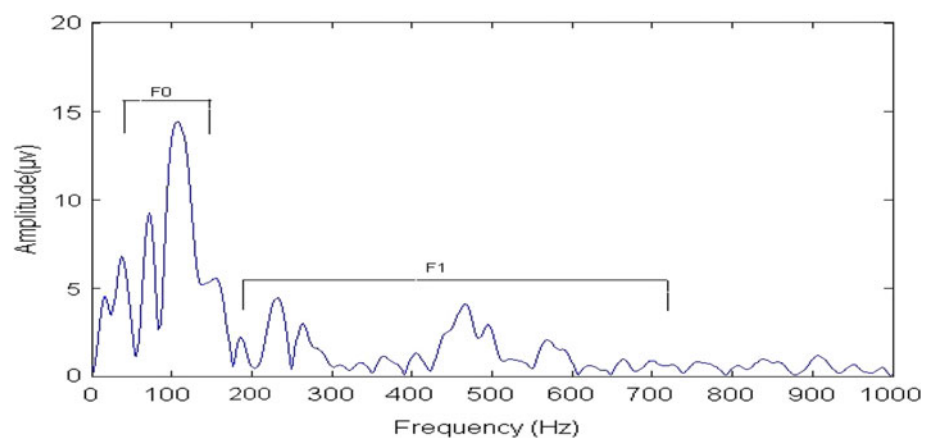
of 6.81 ms. This is more than the documented mean latency of peak V for click stimulus which is reported to range between 5 and 6 ms [15–17]. The latency of the later peaks were calculated with reference to the latency of wavelet V. The positive peak V is analogous to the wave V elicited by a click stimulus. Wave C occurred at a mean latency of 16.82 ms, wave D occurred at a mean latency of 24.75 ms, wave E occurred at a mean latency of 31.36, and wave F occurred at a mean latency of 40.04 ms. The latency obtained for the different peaks of speech evoked ABR in the present study were almost similar to the studies reported earlier [1, 8].

To calculate the amplitude of F0 and F1 fast Fourier analysis was done between 11.5 ms and 46.5 ms of epoch. The measurement of the FFR analysis shows a peak at F0 & F1. It can be seen that the amplitude of the F0 is more compare to the F1. This could be due to the stimulus characteristics and the phase locking properties of the auditory neurons. The stimulus has a higher energy at the F0 region compared to its harmonics [18] and higher energy components are represented better at the neuronal level. Also, the F0 has a lower frequency compared to its harmonics and it is reported that lower the frequency better will be the phase locked responses [19]. Thus the F0 having greater energy and better phase locking is coded more robust than its harmonics.

Conclusions

Speech evoked ABR was present in all the normal hearing subjects here in the present study. The present study highlighted the processing of speech in brainstem. The understanding of speech evoked ABR have other applications both in research as well as in clinical settings. This is specially important if one is interested in studying the central auditory system function. The major application of speech evoked ABR can be in diagnosing and categorizing

Fig. 3 Analysis of F0 and F1. Response indicates that only the fundamental frequency and first formant frequency (F0 = 103–121 Hz; F1 = 220–720 Hz) were measurable



children with learning disability in different subgroups, assessing the effects of aging on central auditory processing of speech, assessing the effects of central auditory deficits in hearing aid and cochlear implant users.

Acknowledgments The Director, All India Institute of Speech and Hearing, Mysore. The students who participated for the study. Dr. Nina Kraus and Dr. Erika Skoe for providing the Brainstem tool Box and the Stimuli used in this study.

References

- Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 115:2021–2030
- Starr A, Don M (1988) Brain potentials evoked by acoustic stimuli. In: Picton TW (ed) *Handbook of electroencephalography and clinical neurophysiology*. Elsevier, Amsterdam, pp 97–150
- Banai K, Nicol T, Zecker S, Kraus N (2005) Brainstem timing: implications for cortical processing and literacy. *J Neurosci* 25:9850–9857
- Galbraith G, Arbagey P, Branski R, Comerci N, Rector P (1995) Intelligible speech encoded in the human brain stem frequency-following response. *NeuroReport* 6(17):2363–2367
- King C, Warrier C, Hayes E, Kraus N (2002) Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neurosci Lett* 319(2):111–115
- Johnson K, Nicol T, Kraus N (2005) Brain stem response to speech: a biological marker of auditory processing. *Ear Hear* 26(5):424–434
- 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. *Biol Psychol* 67(3):299–317
- Johnson K, Nicol T, Zecker S, Kraus N (2008) Developmental plasticity in the human auditory brainstem. *J Neurosci* 28(15):4000–4007
- Krishnan A, Xu Y, Gandour J, Cariani P (2004) Human frequency following responses: representation of pitch contours in Chinese tones. *Hear Res* 189(1–2):1–12
- Song J, Skoe E, Wong PCM, Kraus N (2008) Plasticity in the adult human auditory brainstem following short term linguistic training. *J Cognitive Neurosci* 20(10):1892–1902
- Prucell DW, John SM, Schneider BA, Picton TW (2004) Human temporal acuity as assessed by envelope following responses. *J Acoust Soc Am* 116(6):3581–3593
- Russo N, Nicol TG, Zecker SG, Hayes EA, Kraus N (2005) Auditory training improves neural timing in the human brainstem. *Behav Brain Res* 156(1):95–103
- Sandeep M, Jayaram M (2007) Effect of dichotic stimulus paradigm on Speech Elicited Brainstem response. *J Indian Speech Hearing Assoc* 21:25–29
- Worden FG, Marsh JT (1968) Frequency following (microphonic like) neural responses evoked by sound. *Electroencephal Clin Neurophysiol* 25:42–52
- Beagley HA, Sheldrake JB (1978) Differences in brainstem responses latency with age and sex. *Br J Audiol* 12:69–77
- Jerger J, Hall JW (1980) Effects of age and sex on auditory brainstem responses. *Arch Otorhinolaryngol* 106:387–391
- Jacobson T (1985) *The auditory brainstem responses*. College Hill Press, San Diego
- Ladefoged P (1996) *Elements of acoustic phonetics*. The university of Chicago press, Chicago
- Gelfand SA (1998) Optimizing the reliability of speech recognition scores. *J Speech Lang Hear Res* 41(5):1088–1102