Aud Vestib Res (2019);28(2):75-86.

REVIEW ARTICLE

Speech-evoked auditory brainstem response: a review of stimulation and acquisition parameters

Abdollah Moossavi[1], Yones Lotfi[2], Mohanna Javanbakht[2*], Soghrat Faghihzadeh[3]

[1]- Department of Otolaryngology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[2]- Department of Audiology, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
[3]- Department of Biostatistics and Epidemiology, Zanjan University of Medical Sciences, Zanjan, Iran

Received: 13 Oct 2018, Revised: 8 Dec 2018, Accepted: 15 Dec 2018, Published: 15 Apr 2019

Abstract
Background and Aim: Speech-auditory brainstem response (ABR) as a new test in the field of auditory electrophysiology, examines the auditory processing of stimuli with complex acoustic structures at the subcortical level. In recent years, speech-ABR has been administered to patients with various hearing impairments and people with special auditory skills. Results of these studies are of great interest to researchers in the fields of cognitive and auditory neurosciences. In this study, because of the increasing use of this test, a review of the studies carried out on the origin of this response and the proposed protocols to stimulate, record, and analyze this electrophysiological response are presented.
Recent Findings: The most common stimulus parameters used in the published articles was /da/ stimulus in 40 ms duration and 60-85 dB SPL intensity with the use of alternative polarity and rate of about 10 stimuli per second. The verified and widely-used acquisition parameters include using vertical electrode array with 6000 sweeps and a 30-3000 Hz filtering in a 60-70 ms time window.
Conclusion: In determining the stimulus-record parameters of the speech-ABR test, in addition to considering the necessary minimums, the final values should always be selected based on the objectives and the study group. The unique features of this test for diagnosis and monitoring of auditory processing at supra-threshold levels, calls for comprehensive studies to formulate guidelines for the application of this test in auditory clinics but the basic points mentioned in this paper should be considered in the selection of each parameter.
Keywords: Speech evoked auditory brainstem response; frequency following response; complex sounds

Citation: Moossavi A, Lotfi Y, Javanbakht M, Faghihzadeh S. Speech-evoked auditory brainstem response: a review of stimulation and acquisition parameters. Aud Vestib Res. 2019;28(2):75-86.

* **Corresponding author:** Department of Audiology, University of Social Welfare and Rehabilitation Sciences, Daneshjoo Blvd., Evin, Tehran, 1985713834, Iran. Tel: 009821-22180066
E-mail: m.javanbakht@yahoo.com

Introduction
Electrophysiological studies of the brainstem through recording and review of auditory brainstem response (ABR) have been conducted since the early 1970s and widely used in various areas such as assessment of auditory thresholds estimation and neurological examination at the brainstem level [1]. Among the main reasons for

the clinicalization of this electrophysiologic test, experts point to its non-invasiveness, reproducibility, sensitivity, and specificity in determining some neurological disorders as well as its frequency characteristic with the use of tone-burst stimuli. The stimuli that are used to record the ABR are generally in the form of click or tone-burst. Click stimulus according to its characteristics (transient nature and stimulated frequency range) provides the best recording waveform with acceptable morphology and transient responses at the brainstem level within 1 to 10 ms range after the stimulus, which increases in some cases such as infants. However, two sets of transient and sustained responses can be defined in terms of time dependence on the stimulus at the brainstem level. The first case is evoked by transient stimuli and the second is due to the periodic pattern in the stimuli. Therefore, in the conventional ABR recording with click stimuli, only the transient responses of the auditory brainstem nuclei are recorded. On the other hand, the conventional ABR recording with a click and even tone-burst is unable to test the ability of the hearing system in real-world interactions with speech and musical stimuli that require both transient responses and sustained ones [2].

In 1980, the earliest studies of hearing-related neuroscience were published on the use of complex and speech stimuli in evaluating brainstem responses [3]. These responses showed speech stimuli encoded with the preservation of the initial signal characteristics at the brainstem level, which confirmed by converting the recorded neural signals to the equivalent input audio signals [4]. The exact imitation of the stimuli is also present in early auditory responses, including cochlear microphonic (CM) [5]. Given that brainstem responses, stimulated with speech, have a delay of about 6 ms after providing an acoustic stimulus, we will certainly face the frequency following response (FFR), and the response cannot be CM because it cannot be recorded simultaneously with the trigger [4,6]. These brainstem-evoked responses to speech, musical, or non-speech spectrally and temporally designed stimuli are called complex-auditory brainstem response (C-ABR). Since the focus of many studies was on single-syllabic speech structures (as short time speech stimuli which are capable of inducing transient and sustained auditory responses at the brainstem level) [7-9], the present article generally uses the name "Speech-evoked auditory brainstem response" or simply speech-ABR for this response.

The speech-ABR is an important objective and non-invasive test in identifying the role of brainstem nuclei like inferior colliculus (IC) which is an important nucleus with extensive communication with auditory and non-auditory centers and also the center of neural interactions of the afferent and efferent auditory pathways [10]. Thus many studies have been conducted on neurological applications of this test. The speech-ABR test is an objective test with speech stimuli in different languages [11-13] that allows the examination of the afferent-efferent auditory processes on speech stimuli at different levels of the brainstem. In addition, it can have a special place in the examination of abilities and impairments of auditory processes. People with problems of auditory processing in various disorders (such as autism, learning disorders, auditory processing disorder, and stutter) show abnormalities in this test. On the other end of the spectrum, people with special auditory experience and skills (like musicians, bilinguals and multi-linguals) also show the temporal and or spectral processing advantages over normal people without these special auditory skills. Various studies have used speech-ABR as an electrophysiological test to confirm the results of behavioral tests of speech perception in noise and the role of subcortical structures in speech processing in the presence of noise. Speech-ABR has been recorded with reduced amplitude, increased latency of waves, and poor fundamental frequency (F0) representation in people with poor speech perception in noise [14-16]. The effects of sensorineural hearing loss are more pronounced by increasing latency of initiating and transient parts of the response and lower changes reported in the FFR part [17]. However, patients with (central) auditory processing disorder changes in the onset, the consonant-vowel (CV) transition and the sustained part have suggested the possibility that this test can be used to diagnose and monitor the rehabilitation outcomes of the patients with central auditory processing disorder [18,19]. Other groups that are characterized by a variety of speech and auditory processing disorders in which speech-ABR can detect the differences in their responses compare to the normal group are reading disorders and dyslexia [20,21], stuttering [22,23], and autism patients [24,25]. Musicians, bilinguals, and multilinguals are among the groups with auditory special skills, each has specific characteristics in the auditory response of the brainstem to speech stimuli or non-verbal sophisticated stimuli like music ones. Research has shown better harmonics representation, lower latency of the onset waves, phase locking at high frequencies, higher F0 amplitude, especially in musicians and rehabilitated individuals with musical exercises compared to control groups (the type of music, age, and duration of rehabilitation affect outcomes) [19,26-28]. Better F0 representation and better wave's morphology compare to monolinguals were reported in bilingual and multilingual people, because they need a more accurate representation of F0 to identify and categorize the sounds of an audio object in different languages. These groups show the functional superiority in the presence of noise compared to monolinguals that show stronger cortical to subcortical level connections in bilinguals and multilinguals [29]. It seems that linguistic experiences, through downstream paths affect processes of auditory upward or afferent pathway [30,31]. Based on these properties of speech-ABR and in general complex-ABR in displaying a specific pattern in response to speech stimuli in individuals with special disorders or groups with special auditory experiences, it is referred to the neural signature of each person, which represents the sum of his or her auditory experiences and abilities [32].

Given the potential of the speech-ABR test and lack of generic stimulation and recording guidelines for it, this study reviews the different stimuli and recording parameters used in the published articles in this field. In this study, we try to determine the necessary criteria for selecting the stimulus parameters (such as stimulus type, duration, polarity and rate), the recording parameters (such as the number and method of electrode array, number of sweeps, filtration and time window), and finally the most used one that provide us appropriate morphology and repeatability is expressed.

Methods

In this review article, studies, books and other resources available on the Google Scholar, the US National Library of Medicine and the National Institute of Health (PubMed), Science Direct, Medline, and SID search engines, which are relevant to this research keywords in accordance with the Medical Subject Headings (MeSh) published from 2000 to 2018 (with the exception of the few articles related to the discovery response history) were chosen. Of 355 sources found, in the first step, duplicates were removed (including 19 articles), and after the initial study of the abstracts of 336 remaining items, 60 ones related to the application and stimulation-recording speech-ABR domains and or related to one of the subtitles considered for our review article were chosen. History related resources were added and attempts were made to have a comprehensive look at published studies in the areas related to the stimulation, recording, and interpretation of the speech-ABR test.

Speech-ABR origins

In the response of the brainstem to CV speech stimuli, like /da/, following the conventional ABR waves and after the V peak of speech-ABR, which appears to be in line with the click-ABR peak V [33], a negative trough called A is seen about 6 to 10 ms after the start of the stimulus, indicating the participation of rostral brainstem centers in the response. Thereafter, a negative C wave is seen indicating the consonant to the vowel transmitting portion in the stimulus and the beginning of the sustained or periodic section of the response to the vowel. D, E, and F peaks occur in response to the stimulus periodic section that follows the

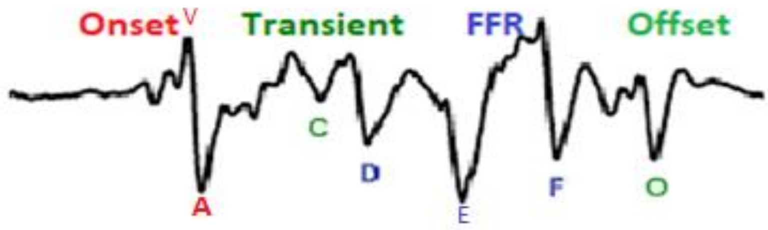

Fig.1. Different parts of the auditory brainstem response evoked by speech /da/ stimulus.

stimulus pattern, and F0 and harmonics are studied with respect to these waves. Finally, the O wave is recorded as a transient response to the offset of the stimulus [34] (Fig. 1).

In various disorders and different test conditions (like silent recording and recording in noise), transient and periodic responses show changes in different manners, so it seems that there are different sources of transient and periodic parts of speech-ABR. For example, recording response in the presence of background noise, mostly affects the latency and amplitude of transient response segments, and the FFR or periodic part of the response is not affected by noise indicating the difference in the function of the centers of transient and periodic responses or, in other words, show the difference in two-part origin of speech-ABR [33-35].

Concerning the origin of the transient and onset parts of speech-ABR, as the latency of the onset section of speech-ABR is about 5 up to 10 ms, it seems that this part of the response originates from brainstem areas like IC [33]. In the case of the periodic section or FFR, due to the latency of this part compared to CM and the disappearance of FFR despite CM preservation in the lesions of the upper brainstem parts like IC, no cochlear origin is suspected for it. On the other hand, because of the much higher latency and amplitude of cortical responses in comparison with the FFR and maintaining the FFR response in sleep conditions despite the removal or reduction of cortical responses, the cortical origin is not considered for the FFR section. This response origin is affected by various centers with phase locking properties at the brainstem level such as cochlear nucleus (CN), superior olivary complex (SOC), and IC. According to Chandrasekaran and Kraus study in 2010, FFR amplitude is mostly related to the role of IC [33]. Regarding transmitted information to the IC through the direct path of the CN to the contralateral IC from the lateral lemniscus (LL) path and the ipsilateral path from CN to SOC, LL, and eventually, the IC, the footprint of other brainstem nuclei can be found in the production of the speech-ABR FFR part. However, because of the missing frequency recording in the vertical electrode array, the importance of rostral brainstem nuclei such as IC and LL would be more prominent in creating the FFR part in speech-ABR, because the contribution of lower brainstem nuclei like CN, generally, is more obvious in the horizontal electrode arrangement so that the missing frequency in the FFR record is not observed in this array [4]. By studying all information about the origin of this response, it seems that IC and rostral levels of brainstem are the most involved parts in producing this response.

In the following overview, first, we will have a general look at the characteristics of the trigger stimulus, the parameters of the record, and then on how to analyze this response.

Stimulus parameters

Each electrophysiological response for general application and the feasibility of comparing the results of various related studies call for a specific guideline and the same stimulus-frameworks, which follows the stimulatory parameters used in various studies. Also, the most used or accepted values for each parameter is mentioned at the end of the relevant section.

According to published studies on the optimal values of some parameters in general use, we still need comprehensive studies that will be mentioned in the relevant sections. Prior to discussing the issues of duration, intensity, polarity, rate, and how to deliver stimulus in a monaural or binaural mode, we will refer to the physical bases and speech-stimulating structure that is capable of stimulating transient and sustained responses at the brainstem level.

Stimulus type

According to published studies in speech-ABR, the least stimulus that can stimulate transient and sustained parts of the brainstem level responses (as well as being the representative of everyday speech) are one-syllable speech stimuli in the form of a CV, such as /da/, /ga/, and /ba/. As with the auditory responses of the brainstem with clicks and tone-bursts stimuli, the faster the rising time and the wider the frequency spectrum of used stimuli will result in the wider activity of the cochlea and the auditory nerve fibers so it can create a better response with greater amplitude and shorter latency. Among the speech consonants, explosive consonants such as /d/, /g/ and /b/ are the best options for the onset transient part of the response. Of course, these three different stimuli due to their differences in F2 and F3 create variations in the final parameters of the response, especially peaks latencies, which indicates the distinction between these stimuli at the level of the brainstem [36]. The most commonly used stimulus is /da/ because the response from the initial transition section of this consonant is similar to the familiar pattern of auditory response to the click stimulus, and its sustained part (the vowel) is also similar to the FFR response that creates with a tonal stimulus. On the other hand, the /da/ syllable is one of the most familiar syllables with similar production in the different languages, which is used in the study of subjects with different languages and showed no significant clinical difference [11-13]. The use of explosion consonants are due to their transient and less energy, which makes it difficult to diagnose and distinguish it in auditory challenging conditions such as speech perception in noise and in patients with auditory processing disorders, which makes the stimulus more difficult [37] and yields more clinical benefits in diagnosing and monitoring the effects of rehabilitation. In the case of FFR, due to the phase locking mechanism in the structures of the brainstem, it is possible to record a response up to a frequency of 1500 Hz, which represents low frequencies and sustained parts of speech and musical stimuli at subcortical levels. In speech production, F0 is the result of the vocal cords vibration during the passage of air between them, because of the asymmetry of the opening and closing of the vocal cords (closing of the cords is much slower than their opening) which is not completely sinusoid and is generally triangular or sawing teeth make frequencies of 80 to 400 Hz along with the harmonics resulting from the resonances generated in the production path. Therefore, in the stimulus design for the recording of speech-ABR response, it is best to use F0 of about 80-300 Hz, which is representative of speech in the real world and in the range of phase locking performance at the brainstem level that allows the acceptable recording of the FFR section [38].

Stimulus duration

Like other electrophysiological responses, the stimulus duration in the speech-ABR test also applies the effects of the stimulus type (consonant, CV, selective vowel), the frequency of the stimulus (low-frequency stimuli need longer stimulation time than high-frequency stimuli), the stimuli intervals, and the response time window. Various studies using CV stimulus lasted 40 ms [39], 170 ms [40], and 350 ms [41], depending on the purpose of the study, the target group, and the time needed to record and analyze the response. Each CV has a unique frequency transition which leads to its identification and differentiation and the vowel's stable section seems to play a minor role in understanding or differentiating [38]. Therefore, in the speech-ABR test, which requires the subject's passive collaboration and reasonable time for clinical application, it is possible to minimize the vowel's sustained time. The duration of about 40 ms can be a selective choice for general use that has been used in many studies, too [11,23,42,43].

Stimulus intensity

Considering that the speech-ABR is designed to investigate the responses of the rostral brainstem in processing speech stimuli, this test is performed at suprathreshold levels (60 to 85 dB SPL). Although the increase in stimulus intensity generally increases the amplitude and decreases the latency due to the increase in the participating neural fibers, the changes are not consistent with the increase in intensity in the transient and sustained part. According to the published reports, changes in stimulus intensity have clearer effects on the sustained section or FFR. For every 10 dB increase, there will be 1.4 ms latency reduction in the sustained part of the response, while the transient part changes in response to the stimulus intensity are more trivial (0.6 dB reduction per 10 dB increase in intensity) [44]. These results suggest different neural processing for the transient and sustained part of the speech-ABR response [38,44].

The usual clinical intensity for /da/ stimulus suggested in the speech-ABR test of evoked potential assessment systems, such as the Intelligence Hearing System (HIS, USA) and Biological Navigator Pro from Natus Medical, USA, is generally about 80 dB SPL. If we use different intensities according to the objectives of the special study, the new norm of the result will be obtained in proportion to the intensity used.

Stimulus polarity

In addition to the polarity changes which occur after transducer's output at different parts of the auditory system and cause ambiguity about the effect of polarity on the final response in all of the auditory evoked responses [5] on the speech stimuli used in speech-ABR studies, it has been shown that the application of various rarefaction, condensation, or alternate polarities does not significantly change the latency of the transient and sustained parts. However, the use of stimulus with the alternating polarity decreases the amplitude of the response, especially in the high-frequency spectral part of the response. On the other hand, considering the advantages of using alternating polarity, such as artifact reduction, noise reduction of the field and also removing CM, it is suggested that the response to individual polarities be recorded in the form of single polarity and then by adding or subtracting them, the best answer is obtained [38,45].

The most common polarity used for speech-ABR at present is alternative one, but it should be noted that in using stimulation with rarefaction or condensation polarity, the norms obtained with alternate polarity or vice versa cannot be used [45]. In general, polarity should be considered depending on the used system and the polarity used for the normalization with that system.

Stimulus presentation rate

The stimulation rate depends on the duration of the stimulus and the time intervals between the stimuli to complete the response to the initial stimulus, return to the baseline, and start the response to the next stimulus. Increasing stimulation rates has the greatest impact on the amplitude and latency of transient responses [5,6] while its effects on the sustained part are minor. This finding again suggests the existence of different neurological resources for the two parts of the response [46]. A distinct difference between the determinations of the stimulation rate in the conventional ABR to the speech-ABR lies in the required stimulus duration.

This time is almost invisible for the click but in the case of speech stimuli, this time is much longer to record the appropriate response and the longer stimulus intervals to prevent the combination of the response of the two stimulants needed in speech-ABR [38]. Thus, the stimulation rate of about 10 stimuli per second is used in general, which like other auditory-evoked responses the numbers of about 10.3, 10.9, or 11.1 have been selected in previous studies, according to the urban frequency (50 or 60 Hz) [11,38,43].

Stimulated ear and binaural interaction component

In general, binaural stimulation with respect to monaural stimulates is more stable and has higher amplitude responses because of binaural accumulation and increased intensity [11,38]. Considering that the speech-ABR test is designed to investigate brainstem functions in the processing of speech stimuli in everyday life and in general, binaural hearing is involved in the process of the auditory signal, it seems that binaural stimulation can better simulate real-life auditory processing. However, in monaural processing, different studies have not reported a significant difference between the left and right ear stimulation, so no noticeable side effect in speech stimulus processing at the brainstem level has been reported [11]. Some others considered slightly shorter latency in right ear stimulation due to the superiority of the right ear in the processing of speech stimuli and emphasized the use of distinct norms for clinical application in stimulating each ear [38,47]. Regarding the recording of the binaural interaction component in the speech-ABR test [48], the recording and comparison of monaural and binaural conditions can have a diagnostic value in binaural interaction disorders, which requires more studies and access to norms in this area.

Stimulus transducer

Providing speech stimulation to record brainstem response is possible through a variety of transducers such as headphones, inserts, and even speakers. However, due to the benefits of insert earphones in increasing interaural attenuation, reducing peripheral noise, and minimizing artifacts, the use of this transducer was the first choice and offer of studies on speech-ABR [26,30,36,42]. Some exceptions should be considered such as hearing aid or cochlear implant users and hard to test groups, which use of insert earphones is impossible and by applying time corrections for the transducer to the ear and doing the necessary calibrations, other alternatives such as speakers or headphones are used depending on the research method [48].

Acquisition parameters

The next question for the application of an electrophysiologic response is how to determine the appropriate acquisition or recording parameters for various purposes and applications. In this section, an overview of the speech-ABR record parameters used in various studies and common values along with the advantages and disadvantages of selecting each of the parameters are discussed.

Electrode montage

The speech-ABR similar to traditional ABRs can be recorded in single- or multi-channel mode. The common method used in single-channel three-electrode recordings is the use of vertical array with conventional placement for negative, positive, and common electrodes on pinna or ear mastoid of the stimulate side, Cz or the upper part of the forehead (FPz) and the pinna or mastoid of contralateral ear, respectively [46,49,50]. The change in each of the electrode positions leads to changes in the morphology, amplitude, and latency of the recorded waves due to the location of the sources that produce this response so it needs to be compared with the normal values obtained with the same arrangement.

State of arousal and attention

Contrary to the conventional ABR recording in sleep and even under anesthesia (in newborns and hard test groups), there is no general agreement between researchers in this area for recording the speech-ABR in sleep. Sleep or anesthesia can reduce muscle artifacts and increase the signal-to-noise ratio. However, according to studies on the effects of auditory experiences and auditory processing capabilities on speech-ABR results, some of these effects on brainstem structure are probably due to the association of cortical centers with subcortical levels through the efferent pathways. Despite the significance and clinical use of speech-ABR to display the mentioned differences in the record of this test in sleep or anesthesia, and reports of some researchers on its recording in sleep [50,51], there is no consensus in this matter. Many studies have performed the test under the inactive subject's co-operation in waking without any attention or performance by the subject, in which they display silent films in order to keep the person awake and make the same test condition in different participants [11,14,52,53]. At the other end of this spectrum, according to fMRI confirmation of the effects of attention on the performance of subcortical structures [54], some speech-ABR studies have been recorded in attentive people [41]. According to the purpose of the study, it seems that different attention situations can be used but in general due to the relatively short duration of this test and the involvement of auditory experiences in responding to speech stimuli, inactive subject co-operation is advisable.

Time window

Given the pre-excitation time to estimate the baseline EEG amplitude, the total response time for the speech stimulus, including the transient, sustained, and consonant to vowel transition sections along with the conventional stimulation time, would be the time span of at least 60 to 70 ms as the preferred option in general studies [23,43,55]. Time window requires adjustments in the event of a change in the other parameters of the stimulus.

Sweeps number

The main criterion for determining the number of sweeps for all the evoked responses is to achieve a proper signal-to-noise ratio, and even for the conventional ABR a specific amount is not specified [5]. Sweep number is variable depending on the response amplitude, ambient noise level, physiological noise amplitude, and other subject and test conditions.

In the studies of Northwestern University as one of the important centers for speech-ABR evaluation in various test groups, 4000 to 6000 sets of sweeps were proposed to obtain a suitable signal-to-noise ratio. In order to optimize the signal-to-noise ratio, adding two 3000-sweep records with identical recording conditions for achieving the final response curve of 6000 sweeps was considered [38].

Filtering

Another approach to improve the signal-to-noise ratio and obtaining optimal response without any artifacts of other centers is the use of appropriate filtering. In the case of speech-ABR for generating high-frequency transient peaks with low-frequency parts of the response, similar to the conventional click-ABR, 100-3000 Hz [38] filter or in conditions that are more conservative, 30-3000 Hz filter is used [56]. However, in a few studies, depending on the used stimulus, the study method and goals, other filter settings such as 0.1 to 1000 Hz have also been used [57], but currently, they have less prevalent in the clinical settings.

Analysis and interpretation of speech-ABR

As mentioned before, brainstem auditory responses to speech stimuli have transient and sustained parts associated with stimuli that proved a stimulus representation in brainstem structures. A response to the common speech stimulus such as /da/, consists of the initial V peak and the first trough A. Then the trough C in the frequency transition region of the stimulus followed by the FFR peaks, including peaks or troughs D, E and F. Ultimately the O peak appears that is transient and refers to offset response (Fig. 1).

In examining the transient peaks associated with the onset, sustained and offset sections, we use common criteria such as amplitude and latency. After accurate determination of the peaks, in addition to studies on the absolute amplitude and latencies of the waves, it is possible to consider the peak-to-peak criteria. The most widely-used criteria in clinical use are the absolute latency and amplitude of the waves plus the slope between the V and A peak, which has shown its diagnostic value in various disorders such as learning disorders [23,58,59].

In the case of the sustained section, methods such as determination of root mean square (RMS), cross-correlation, and auto-correlation are used to examine the overall amplitude, to determine the signal-to-noise ratio, to determine the overall morphology and the similarity of the response to stimulation signal, and to determine

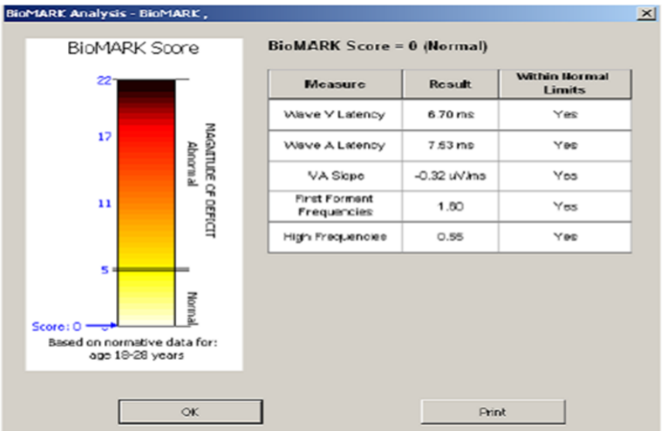

Fig. 2. A sample of scoring and reporting form of speech-ABR recorded by an auditory evoked potential system (Biologic-Navigator Pro, Natus, USA) the AEP system.

the repeat patterns such as the F0 and time envelope, respectively. Another method used in the frequency domain is the Fourier analysis to verify the accuracy and amplitude of the neural phase dependence at a specific frequency or at a specific frequency set. The basis for using fast Fourier transform (FFT) is determining the frequencies in a complex stimulus and the amplitude of each at each time point of the final response. Many evaluations of this section are used outside of the auditory evoked potential (AEP) system and are transmitted to analytical software such as MATLAB, using code written to process a response to a defined stimulus.

In clinical applications for the normal interpretation or abnormal response obtained in speech-ABR, at least 5 criteria, including V wave latency, A wave latency, VA slope, mean amplitude of first formant frequencies and mean response at higher frequencies (proportional to the stimulus) are determined. Then the location of each of these five criteria within the normal range or outside of this range is determined, separately. Finally, each of these criteria is awarded a score that generally represents the response in the form of a number, and this number is compared with the range of normal values for that particular age group (Fig. 2). This is the pre-determined method used for speech-ABR analysis in the Biologic-Navigator Pro system (Natus Medical Inc.) and this analysis uses its own system stimulation and recording conditions. In the event of a change in stimulus or in each of the recording conditions, it is not possible to compare the results with the normal values of the system and the transient and sustained response evaluations should be performed separately and compared with the normal values appropriate to the stimulus used and the subject characteristics.

About the characteristics of the subject, the factors such as age [60], hearing loss [17], and even gender [61,62] or race [49], which can cause significant differences in results should be considered.

Conclusion

The speech-ABR test as a part of auditory-evoked responses to complex stimuli is a new and uncharted area among auditory electrophysiological tests. In recent years, due to the unique ability of this test to display auditory processes and plasticity changes at the brainstem level, it has been of great interest to researchers in the fields of speech-auditory neuroscience and cognitive sciences. Because of the response complexity of the processing, especially the sustained part or its FFR part, a comprehensive clinical application has not been reported yet. However, the results of studies conducted in groups with special auditory skills and groups with auditory processing impairments have been proven the ability of this test as a unique neural signature for displaying the auditory experiences and capabilities.

With regard to these functional capabilities, the guidelines for determining the correct stimulation and recording parameters and key features of speech-ABR analysis were pointed out in this review article. However, as mentioned before, it seems that for general use of this test and the universality of this assessment method and its response analysis, more studies with assessment of more subjects in different societies are necessary to provide general guidelines for this test application like other clinical electrophysiological tests.

Acknowledgments

This article is a part of the M. Javanbakht's Ph.D. dissertation titled "A survey on rostral part of the auditory efferent system based on speech-ABR in noise test" submitted in the Department of Audiology, University of Social Welfare and Rehabilitation Sciences, Tehran.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Jewett DL, Romano MN, Williston JS. Human auditory evoked potentials: possible brain stem components detected on the scalp. Science. 1970;167(3924):1517-8.
2. Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. J Neurosci. 2008;28(15):4000-7. doi: 10.1523/JNEUROSCI.0012-08.2008
3. Greenberg S. Temporal neural coding of pitch and vowel quality. [dissertation]. Los Angeles: University of California; 1980.
4. Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM. Intelligible speech encoded in the human brain stem frequency-following response. Neuroreport. 1995;6(17):2363-7.
5. Hall JW. New handbook of auditory evoked responses. Boston: Pearson Education, Inc; 2007. p. 110-21.
6. Krishnan A. Frequency-following response. In: Burkard RF, Don M, Eggermont JJ, editors. Auditory evoked potentials: basic principles and clinical application. 1st ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 313-29.
7. Wible B, Nicol T, Kraus N. Correlation between brainstem and cortical auditory processes in normal and language-impaired children. Brain. 2005;128(2):417-23. doi: 10.1093/brain/awh367
8. Kraus N, Nicol T. Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends in neurosciences. 2005;28(4):176-81. doi: 10.1016/j.tins.2005.02.003
9. Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N. Auditory training improves neural timing in the human brainstem. Behavioural brain research. 2005; 156(1):95-103 doi: 10.1016/j.bbr.2004.05.012
10. Stebbings KA, Lesicko AM, Llano DA. The auditory corticocollicular system: molecular and circuit-level considerations. Hear res. 2014;314:51-9. doi: 10.1016/j.heares.2014.05.004
11. Ahadi M, Pourbakht A, Jafari AH, Jalaie S. Effects of stimulus presentation mode and subcortical laterality in speech-evoked auditory brainstem responses. Int J Audiol. 2014;53(4):243-9. doi: 10.3109/14992027.2013.866281
12. Rana B, Barman A. Correlation between speech-evoked auditory brainstem responses and transient evoked otoacoustic emissions. J Laryngol Otol. 2011;125(9):911-6. doi: 10.1017/S0022215111001241
13. Kouni SN, Giannopoulos S, Ziavra N, Koutsojannis C. Brainstem auditory evoked potentials with the use of acoustic clicks and complex verbal sounds in young adults with learning disabilities. Am J Otolaryngol. 2013;34(6):646-51. doi: 10.1016/j.amjoto.2013.07.004
14. Anderson S, Parbery-Clark A, Yi H-G, Kraus N. A neural basis of speech-in-noise perception in older adults. Ear Hear. 2011;32(6):750-7. doi: 10.1097/AUD.0b013e31822229d3
15. Maruthy S, Kumar UA, Gnanateja GN. Functional interplay between the putative measures of rostral and caudal efferent regulation of speech perception in noise. J Assoc Res Otolaryngol. 2017;18(4):635-48. doi: 10.1007/s10162-017-0623-y
16. Song JH, Skoe E, Banai K, Kraus N. Perception of speech in noise: neural correlates. J Cogn Neurosci. 2011;23(9):2268-79. doi: 10.1162/jocn.2010.21556
17. Nada NM, Kolkaila EA, Gabr TA, El-Mahallawi TH. Speech auditory brainstem response audiometry in adults with sensorineural hearing loss. Egyptian journal of ear, nose, throat and allied sciences. 2016;17(2):87-94. doi: 10.1016/j.ejenta.2016.04.002
18. Rocha-Muniz CN, Mamede Carvallo RM, Schochat E. Medial olivocochlear function in children with poor speech-in-noise performance and language disorder. Int J Pediatr Otorhinolaryngol. 2017;96:116-21. doi: 10.1016/j.ijporl.2017.03.003
19. Martin JM, Billiet CR, Bellis TJ. Audiologic Assessment of CAPD. In: Geffner D, Ross-Swain D, editors. Auditory processing disorders: assessment, management, and treatment. 3rd ed. San Diego: Plural Publishing; 2018. p. 69-94.
20. Kouni SN, Papadeas ES, Varakis IN, Kouvelas HD, Koutsojannis CM. Auditory brainstem responses in dyslexia: comparison between acoustic click and verbal stimulus events. J Otolaryngol. 2006;35(5):305-9.
21. Skoe E, Brody L, Theodore RM. Reading ability reflects individual differences in auditory brainstem function, even into adulthood. Brain Lang. 2017; 164:25-31. doi: 10.1016/j.bandl.2016.09.003
22. Gonçalves IC, De Andrade CRF, Matas CG. Auditory processing of speech and non-speech stimuli in

children who stutter: electrophysiological evidences. Brain Disord Ther. 2015;4(5):199.
23. Tahaei AA, Ashayeri H, Pourbakht A, Kamali M. Speech evoked auditory brainstem response in stuttering. Scientifica (Cairo). 2014;2014:328646. doi: 10.1155/2014/328646
24. Russo N, Zecker S, Trommer B, Chen J, Kraus N. Effects of background noise on cortical encoding of speech in autism spectrum disorders. J Autism Dev Disord. 2009;39(8):1185-96. doi: 10.1007/s10803-009-0737-0
25. Russo N, Nicol T, Trommer B, Zecker S, Kraus N. Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Dev Sci. 2009;12(4):557-67. doi: 10.1111/j.1467-7687.2008.00790.x
26. Skoe E, Kraus N. Musical training heightens auditory brainstem function during sensitive periods in development. Front Psychol. 2013;4:622. doi: 10.3389/fpsyg.2013.00622
27. Kumar P, Anil SP, Grover V, Sanju HK, Sinha S. Cortical and subcortical processing of short duration speech stimuli in trained rock musicians: a pilot study. Eur Arch Otorhinolaryngol. 2017;274(2):1153-60. doi: 10.1007/s00405-016-4285-x
28. Strait DL, Kraus N. Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. Hear Res. 2014;308:109-21. doi: 10.1016/j.heares.2013.08.004
29. Lotfi Y, Chupani J, Javanbakht M, Bakhshi E. Evaluation of speech perception in noise in Kurd-Persian bilinguals. Aud Vestib Res. 2019;28(1):36-41. doi: 10.18502/avr.v28i1.414
30. Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci U S A. 2012;109(20):7877-81. doi: 10.1073/pnas.1201575109
31. Krizman J, Skoe E, Marian V, Kraus N. Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. Brain Lang. 2014;128(1):34-40. doi: 10.1016/j.bandl.2013.11.006
32. Kraus N, Anderson S. Identifying neural signatures of auditory function. Hear J. 2015;68(1):38-40. doi: 10.1097/01.HJ.0000459742.94251.cc
33. Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology. 2010;47(2):236-46. doi: 10.1111/j.1469-8986.2009.00928.x
34. Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clin Neurophysiol. 2004;115(9):2021-30. doi: 10.1016/j.clinph.2004.04.003
35. Cunningham J, Nicol T, King C, Zecker SG, Kraus N. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. Hearing Research. 2002;169(1-2):97-111. doi: 10.1016/S0378-5955(02)00344-1
36. Johnson KL, Nicol T, Zecker SG, Bradlow AR, Skoe E, Kraus N. Brainstem encoding of voiced consonant--vowel stop syllables. Clin Neurophysiol. 2008;119(11):2623-35. doi: 10.1016/j.clinph.2008.07.277
37. Lotfi Y, Kargar S, Javanbakht M, Biglarian A. Development, validity and reliability of the Persian version of the consonant-vowel in white noise test. Journal of Rehabilitation Sciences and Research. 2016;3(2):29-34
38. Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear Hear. 2010;31(3):302-24. doi: 10.1097/AUD.0b013e3181cdb272
39. Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. Cereb Cortex. 2009;19(11):2699-707. doi: 10.1093/cercor/bhp024
40. Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. J Neurosci. 2012;32(41):14156-64. doi: 10.1523/JNEUROSCI.2176-12.2012
41. Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci U S A. 2007;104(40):15894-8. doi: 10.1073/pnas.0701498104
42. Sinha SK, Basavaraj V. Speech evoked auditory brainstem responses: a new tool to study brainstem encoding of speech sounds. Indian J Otolaryngol Head Neck Surg. 2010;62(4):395-9. doi: 10.1007/s12070-010-0100-y
43. Ansari MS, Rangasayee R, Ansari MA. Neurophysiological aspects of brainstem processing of speech stimuli in audiometric-normal geriatric population. J Laryngol Otol. 2017;131(3):239-244. doi: 10.1017/S0022215116009841
44. Akhoun I, Gallégo S, Moulin A, Ménard M, Veuillet E, Berger-Vachon C, et al. The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. Clin Neurophysiol. 2008;119(4):922-33. doi: 10.1016/j.clinph.2007.12.010
45. Kumar K, Bhat JS, D'Costa PE, Srivastava M, Kalaiah MK. Effect of stimulus polarity on speech evoked auditory brainstem response. Audiol Res. 2014;3(1):e8. doi: 10.4081/audiores.2013.e8
46. Krizman JL, Skoe E, Kraus N. Stimulus rate and subcortical auditory processing of speech. Audiol Neurootol. 2010;15(5):332-42. doi: 10.1159/000289572
47. Hornickel J, Skoe E, Kraus N. Subcortical laterality of speech encoding. Audiol Neurootol. 2009;14(3):198-207. doi: 10.1159/000188533
48. Dabbous AO, Koura RA, Hamdy MM. Complex auditory brainstem response in normal-hearing adults using binaural versus monaural speech stimuli. The Egyptian Journal of Otolaryngology. 2017;33(4):656-62. doi: 10.4103/ejo.ejo_74_16
49. Song JH, Skoe E, Banai K, Kraus N. Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex. 2012;22(5):1180-90. doi: 10.1093/cercor/bhr196
50. Sanguebuche TR, Peixe BP, Bruno RS, Biaggio EPV, Garcia MV. Speech-evoked brainstem auditory responses and auditory processing skills: a correlation in adults with hearing loss. Int Arch Otorhinolaryngol. 2018;22(1):38-44. doi: 10.1055/s-0037-1603109
51. Dajani HR, Purcell D, Wong W, Kunov H, Picton TW. Recording human evoked potentials that follow

the pitch contour of a natural vowel. IEEE Trans Biomed Eng. 2005;52(9):1614-8.. doi: 10.1109/TBME.2005.851499
52. Song JH, Nicol T, Kraus N. Test-retest reliability of the speech-evoked auditory brainstem response. Clin Neurophysiol. 2011;122(2):346-55. doi: 10.1016/j.clinph.2010.07.009
53. Heidari A, Moossavi A, Yadegari F, Bakhshi E, Ahadi M. Effects of age on speech-in-noise identification: subjective ratings of hearing difficulties and encoding of fundamental frequency in older adults. J Audiol Otol. 2018;22(3):134-139. doi: 10.7874/jao.2017.00304
54. Rinne T, Balk MH, Koistinen S, Autti T, Alho K, Sams M. Auditory selective attention modulates activation of human inferior colliculus. J Neurophysiol. 2008;100(6):3323-7. doi: 10.1152/jn.90607.2008
55. Anderson S, Kraus N. Neural encoding of speech and music: implications for hearing speech in noise. Semin Hear. 2011;32(2):129-141. doi: 10.1055/s-0031-1277234
56. Galbraith GC, Amaya EM, de Rivera JM, Donan NM, Duong MT, Hsu JN, et al. Brain stem evoked response to forward and reversed speech in humans. Neuroreport. 2004;15(13):2057-60.
57. Reichenbach CS, Braiman C, Schiff ND, Hudspeth AJ, Reichenbach T. The auditory-brainstem response to continuous, non-repetitive speech is modulated by the speech envelope and reflects speech processing. Front Comput Neurosci. 2016;10:47. doi: 10.3389/fncom.2016.00047
58. Ramezani M, Lotfi Y, Moossavi A, Bakhshi E. Auditory brainstem response to speech in children with high functional autism spectrum disorder. Neurol Sci. 2018. doi: 10.1007/s10072-018-3594-9
59. Malayeri S, Lotfi Y, Moossavi SA, Rostami R, Faghihzadeh S. Brainstem response to speech and non-speech stimuli in children with learning problems. Hear Res. 2014;313:75-82. doi: 10.1016/j.heares.2014.04.009
60. Vander Werff KR, Burns KS. Brain stem responses to speech in younger and older adults. Ear Hear. 2011;32(2):168-80. doi: 10.1097/AUD.0b013e3181f534b5
61. Ahadi M, Pourbakht A, Jafari AH, Shirjian Z, Jafarpisheh AS. Gender disparity in subcortical encoding of binaurally presented speech stimuli: An auditory evoked potentials study. Auris Nasus Larynx. 2014;41(3):239-43. doi: 10.1016/j.anl.2013.10.010
62. Jalaei B, Azmi MHAM, Zakaria MN. Gender differences in binaural speech-evoked auditory brainstem response: are they clinically significant? Braz J Otorhinolaryngol. 2018. doi: 10.1016/j.bjorl.2018.04.005