REVIEW ARTICLE

Open Access

Speech auditory brainstem responses (s-ABRs) as a new approach for the assessment of speech sounds coding

Hagar Ahmed Elsayed^{1*}, Ebtessam Hamed Nada¹, Dalia Helal Galhoum¹ and Nadia Mohamed Elnabtity¹

Abstract

Background The evoked auditory brainstem response (EABR) is an objective electrophysiological test used to assess the brainstem's auditory neural activity. Speech ABR (s-ABR) testing using verbal stimuli gives more important details about how the brainstem processes speech inputs which enables the detection of auditory processing impairments that do not manifest in click-provoked ABR. The use of speech syllables in the s-ABR reveals an important brainstem function that plays a crucial part in reading development and phonologic achievement which is an assessment of speech syllables. The syllable /da/ is often utilized in s-ABR measurement being a global syllable that can be tested in many nations with effective experimental confidence.

Conclusion The speech ABR is an objective, unbiased, quick test. It can be used to differentiate between many conditions such as auditory processing disorders (APD), specific language impairment (SLI), and children with academic challenges.

Keywords Speech syllable, Electrophysiology, s-ABR, Central auditory processing, SLI

Background

Auditory evoked potentials (AEPs) are objective measurements of electric discharge in the auditory pathway in response to sounds. Responses are taken from multiple stations from the eighth cranial nerve and up to the auditory cortex. Following the start of the auditory stimuli, different auditory system regions produce their responses at various times which are called latencies (Fig. 1) [1].

The early AEP includes the auditory brainstem response (ABR) whose waveform is composed of five to seven peaks designated as I, II, III, IV, V, VI, and VII [2] which can be detected early in life. The first detected are waves I, III, and V. They appear in infancy followed by the adult-like ABR morphology with

Hagar Ahmed Elsayed

hagarahmedoo660@gmail.com

Zagazig University, Zagazig, Egypt

mature latency and amplitude standards detected at roughly 2 years of age [3].

Higher centers produce later AEP responses, later is the auditory late response (ALR) which has only the (P1) peak of the visible during infancy, while the remaining responses (P1, N1, P2, N2) attain growth and have shorter latencies till reaching adulthood between the ages of 16 and 18 years (Fig. 1) [1].

The auditory middle latency response (AMLR) waveform responses are obtained from centers in the middle of the auditory pathway and labeled Na, Pa, Nb, and Pb. They grow to mature form, to become similar to those of adults concerning latencies and amplitude at approximately the age of 10 years with only the (Na) peak visible throughout infancy [1].

Varied states of arousal have different effects on different types of AEP. For instance, ALR can only be performed on awake and attentive subjects, and AMLRs are impacted by the acting of sleep and are not measurable in sedated persons [1].



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.



^{*}Correspondence:

¹ Audio-Vestibular Medicine Unit, ENT Department, Faculty of Medicine,



Fig. 1 Illustration of AEP waveforms and their peaks showing the ABR peaks (I, II, III, IV, V, and VI) occurring within the first 7 ms, followed by the AMLR (labeled MLR in this figure) peaks (Na, Pa, Nb, Pb [also ALR P1]) and followed by the ALR (labeled LLR in this figure) peaks (P1, N1, P2, and N2) (reprinted from Khuwaja et al., 2015) [1]

It has been observed that when someone is listening attentively to the sound being received, the N1 and P2 responses amplify, especially for low-intensity readings of the stimulus [4]. Similarly, the P300 which is one of the ALR responses that arises more than 300 ms after the N2 peak can only be identified by someone who is listening intently to the stimuli [5].

On the other hand, ABRs are unaffected by the state of arousal and could be obtained during either normal or induced sleep and even under general anesthesia [6].

As regards stimuli used to evoke ABR, in addition to click, tone-burst, and chirp, stunted consonant–vowel (CV) spoken syllables like [ba], [da], and [ga] have been used [6].

s-ABR is a relatively new distinctive procedure to obtain brainstem response since it may be used to quantify the subcortical encoding of speech, and it indicates the acoustic features of the stimulus that evoked it. Researchers currently use it to assess subcortical recognition of consonants and vowels (CVs) as well as comparing subcortical storage of speech in noise to behavioral results of speech detection in noise in normal hearing individuals [7]

The s-ABR measurement is repeatable within and between sessions in normal-hearing infants and adults with all waveform components detected from infancy to older age [8]. The same stimuli were also utilized to record innate speakers of other languages, such as Arabic, Hebrew, and Indian, using their s-ABRs. Speech ABR is a potential technology for clinical use due to the consistency and detection of all response elements among age groups and languages [9]. As illustrated in Fig. 2 [7], the s-ABR contains an onset, transition, frequency following response (FFR), and offset response in reaction to a brief CV (e.g., 40 ms [da]). As seen in Fig. 3, the s-ABR as a response to a prolonged CV (such as 170 ms [da]) shows a further persistent FFR [10].

First, a response with a positive peak V, like the peak V of the click ABR, and a negative trough after peak V, known as peak A, is found about 6 to 10 ms following the stimulus' onset. This suggests that rostral brainstem centers were immersed in the production of the response of the s-ABR [4]. This onset response has an identical neuronal originator like those of wave V of the click ABR and is elicited by the beginning of the consonant in the CVs stimuli (beginning of sound) [11]. Since the speech ABR's onset component has a latency of between 5 and 10 ms, it appears that the brainstem regions such as the inferior colliculus (IC) are the cause of this part of the response.

Second is the transition response, which has negative troughs B and C that are induced by the change in the stimulus from consonant to vowel, but these are not always present in all people [11].

Next are the negative peaks D, E, and F in the frequency following response (FFR), generated by the vowel of the short CV or the vowel formant transition period of the longer CV. There is also the sustained FFR produced in response to the longer CV. It contains periodic peaks that are phase-locked to F0 of the sustained vowel of the CV stimulus (Fig. 4) [10].

FFR is of brainstem origin, suggested by its disappearance in cases of upper brainstem injuries like IC lesions and, moreover, by comparison of its latency (5-10 ms) to the cochlear microphonics (1-2 ms). Its





Fig. 2 Forty-millisecond [da] stimulus (top) and speech ABR in response to the 40 ms [da] (bottom) [9]



Fig. 3 One hundred seventy-millisecond [da] stimulus (top) and speech ABR in response to the 170 ms [da] (bottom) [9]

cerebral origin is also excluded by the FFR's longer latency and the possibility that it was recorded while the subject was sleeping [13].

As a result, it is thought that the cochlear nucleus, superior olivary complex, lateral lemniscus, and inferior colliculus are some of the brainstem regions from which



Fig. 4 Speech-evoked auditory brainstem response [12]

FFR is produced with the main generator being the inferior colliculus [10]. The top-down processing and corticofugal descending ways are moreover immersed in the FFR part of the s-ABR.

Finally, the offset response involves a negative trough (O) induced by the finish of the sound (sound offset) [7].

The CV stimulation pattern that causes the s-ABR is similar to this waveform. Both briefer and lengthier CVs have been employed in the s-ABR information. In addition to an initial burst, a vowel formant transition phase, a persistent vowel period, and shortened CVs (such as 40 ms [da]) also exhibit these features [14]. A persistent vowel period, a formant transition phase, and an onset burst are all present in the lengthier CVs (such as 170 ms [da]) [15].

The s-ABR also includes extra CV stimulus elements listed below: (I) the length of the vowel's F0 reflects the latency of the tiny troughs between the peaks D, E, and F; (II) the length of the vowel's F1 reflects the latency of the little troughs between the peaks D, E, and F; and (III) even if the vowel's frequencies exceed the brainstem's ability to lock it in phase, the vowel's F2 affects speech ABR peak latencies [7, 15, 16].

When compared to CVs with lower F2 frequencies, CVs with higher F2 frequencies create s-ABRs more quickly. Pitch is defined as having the lowest frequency, or F0; formants, such as F1 and F2, are a collection of harmonics, or element frequencies, that are multiples of F0. Nevertheless, each vowel has three formants (F1, F2, and F3) that are unique to that vowel and are used to identify vowels [16].

Additional neurological components of the s-ABR were hypothesized by Kraus and Nicol (2005). According to their idea, auditory cortical circuits start in the brainstem and are shown in the s-ABR. These auditory pathways are the two sensory pathways that the auditory cortex is thought to use to identify the speaker, identify sounds, and determine where sounds are coming from

[16]. They contend that the s-ABR peaks D, E, F, and F0 serve as a representation of the pathways by indicating the nonlinguistic characteristics of the signal that aid in speaker identification. Peaks V, A, C, and O represent the pathways because they show how the articulators move to create speech that shows "where in frequency" the sounds are present [17].

When it comes to speech recognition and communication, the s-ABR components are useful because the start of sound is crucial for identifying phonemes. Frequency changes are essential for identifying consonants and identifying suprasegmental speech characteristics, vowel identification depends on formant structure, and F0 includes non-linguistic data-like emotion and gender [16, 18]. As a result, the s-ABR is capable of measuring objectively essential auditory cues for speech recognition.

On the other hand, to test CV discrimination, the s-ABR can also be measured using CVs with different vowel F2 frequencies. Despite exceeding the top threshold for brainstem phase-locking to sound stimuli, the F2 frequency (> 1000 Hz) is yet exemplified in the brainstem response in conditions of response timing [19].

Three 170-ms CVs with varying F2 frequencies were used to elicit s-ABR trough latencies in children, and these latencies varied. Peak latencies followed [ga] responses earlier than [da] and [ba] responses later than [da]. The highest F2 frequency was found in the [ga], then [da] and [ba], which had the lowest possible F2 frequency [15].

The phase timing analysis revealed that in both children and adult musicians, the phase of the s-ABR in response to [ga] led the phase in response to [da], and the phase in response to [da] led the phase in response to [ba]. Therefore, the speech ABR can serve as a trustworthy audiological marker of subcortical auditory difference [20].

In addition, the peaks V, A, D, and F of the aided s-ABRs in both credentials were more than those of the

unaided s-ABRs. Additionally, in both contexts, there are greater peaks in aided s-ABRs than in unaided speech. These findings are to be expected given that aiding makes sounds louder and hence more audible, which causes earlier latencies, larger amplitudes, and improved response identification [9].

Although it was expected that amplitude would grow as sound level and audibility increased, there was no discernible difference in the amplitude of troughs E and O. Peak E is one of the three FFR troughs elicited by the vowel of the [da]; therefore, it is puzzling why variations in FFR amplitudes were not noticed among the three peaks (D, E, and F). The literature has not yet discussed the connections between these three peaks and stimulus intensity. Additionally, specific effects of stimulus level on the offset peak O have not been observed; nonetheless, the addition of background noise does not appear to have an impact on peak O amplitude. This is possibly a consequence of a compensation mechanism within the brainstem pathway that was reported by Russo et al. [21] This compensation may be the reason why the amplitude of peak O was not affected by a change in audibility [9].

Thus, s-ABR assesses speech recognition, CV stimulus discrimination, and speech-in-noise presentation both with and without hearing aids and can potentially be used as an unbiased assessment of the benefits of HAs for those with sensorineural hearing loss (SNHL) [15].

The best ways to enhance auditory function, particularly the process of voice perception, are through auditory training and amplification. s-ABR testing can show the neurophysiological changes that auditory training can bring about. The auditory training program encourages improvements in speech perception in both calm and noisy situations, as well as in short-term memory and attention processes [22].

In this circumstance, the evaluation of s-ABR may play a significant role in outlining the true benefits of intervention in an objective manner. To select participants who will benefit from an auditory training program, the assessment of s-ABR is therefore thought of as a clinical marker of auditory training [23].

The efficacy of the auditory training program may be determined by the s-ABR evaluation. To ascertain whether this kind of evaluation can be useful in monitoring the older population, more research is required [23].

Although ALRs have been effectively used as objective outcome measures in adults and children with CIs, they have not yet become a part of common clinical practice. This may be due to a variety of physiological features of ALRs, including (i) the fact that ALRs are not established and resemble adults until between the ages of 16 and 18; (ii) the fact that attention and state of arousal have an impact on ALRs, with the response being more pronounced when the subject is paying attention to the provocation as opposed to when they are not; and (iii) ALRs are affected by sleeping and anesthesia [3].

Auditory brainstem responses (ABR) have some rewards over ALRs, including (i) early maturation; (ii) accurate measurement in infants, children, and adults with special requirements; (iii) greater consistency than ALRs within and across subjects; and (iv) being not dependent on attention, state of arousal, sedation, and anesthesia [3].

Consequently, s-ABR can additionally be used as a clinical indicator of the value of CI. However, there is little research on s-ABRs in CI recipients. This is probably because the CI produces a significant number of electrical artifacts, which could obscure the s-ABR waveform [9].

While s-ABR seems to be stable by the age of five, the response to nonverbal stimulus in ABR takes about 18 months to reach maturity [15]. As a result, a technique might be used on younger children and adolescents in school to help differentiate between illnesses that have similar symptoms. The effect of an individual's age on how sounds are coded by a single stimulus or a complex with enhanced hearing ability and brain timing is examined in order to confirm the age at which the central auditory system matures for speech sounds and to determine typical values for various age groups. It was found that, when compared to children between the ages of 8 and 12, a child's s-ABR responses at age 5 are not significantly different, while those at ages 3 to 4 show significant morphological differences concerning latency time [24].

Gender has an effect on s-ABR morphology where women had greater responses (higher amplitudes and lower latency values) than men, and this may be attributed to the effect of estrogen activity [25].

The neuronal synchronization is decreased in the elderly which is presented with problems in the encoding of speech sounds, especially when the speech is delivered in noise. The elderly reported difficulty in understanding speech in noise can be evaluated using the s-ABR assessment. Using hearing aids permits speech to be perceived more obviously, and so, there has been a change in waveform and latency assessments of the aided speech responses ABR [26]. Thus, s-ABR can be used to evaluate the impact of central auditory abilities and the effects of aging on speech processing in the brain [27].

Long-term intensive musical training seems to change anatomy and physiology, as well as enhance working memory in cognitive procedures, emotion control, and auditory perception. The brainstem plays an essential function in determining speech stimuli and temporal feature processing [27]. The perception of consonant duration, also the recognition of notes, and musical scales are both influenced by temporal processing. Temporal processing has an impact on all aspects of the literacy process, involving language, reading, and writing. The identification of minute and quick alterations in sound stimulation is related to rhythm, frequency, phonemic discrimination, duration, and pitch discrimination. To comprehend how music affects the storage of speech sounds and the learning process, analysis of the responses of speech ABR is therefore helpful [27].

s-ABRs can be measured with background noise to assess how noise affects the response. It is known that adding noise may alter the s-ABR waveform in ways that include (i) delaying trough latencies, (ii) reducing peak amplitudes, (iii) decreasing F0 amplitudes, and (iv) decreasing the accuracy and reliability with which the global response accurately and reliably reflects the spectral and temporal properties of the inducement [28]. The s-ABR peaks V and A (onset peaks), the FFR period induced by the vowel formant transitions (peaks D, E, and F), and F0 have both been demonstrated to be more affected by background noise than the sustained FFR period evoked by the steady-state vowel. It has been demonstrated that these modifications to s-ABRs are stable, repeatable, and applicable to different test sessions and participants [29].

The effect of the background noise of s-ABR was evident as presented by the deterioration of latencies (prolongation) and amplitudes (smaller) in persons who scored poorer in behavioral speech-in-noise assessment tests than in those who scored higher in similar behavioral speech-in-noise trials. Consequently, the s-ABR can be utilized as a clinically accurate indicator of how well speech is performed in noisy environments [30]. It can also be used to measure efferent activity at the brainstem level and examine the rostral (top) auditory efferent system's performance because it is more active in noisy environments [31].



Fig. 5 Effects of aiding: displayed showing earlier latencies and larger amplitudes in the aided compared to unaided speech ABRs in quiet (a) and in noise (b). Effects of background noise: displayed limited effects of noise on both aided (c) and unaided (d) speech ABR latencies and amplitudes [9]

In addition to suffering from background noise and competing sounds, children with learning, speech, and hearing impairments also have some difficulties perceiving speech sounds in quiet settings. This issue may result from modifications in temporal processing that affect how speech is perceived. In this situation, the s-ABR can be utilized to identify children who are predisposed to these changes and serve as a clinical indicator of auditory processing problems [33].

Children with dyslexia frequently experience difficulties perceiving speech sounds, which may impair their reading abilities. Those with unreliable neural reactions show some difficulty when learning to read. Good readers have an accurate neural representation of sound. Thus, the s-ABR can aid in classifying and identifying these children so that a more suitable involvement can be provided. The s-ABR can be used to identify and classify distinct subgroups of children with learning disabilities [34].

Conclusion

Proper functioning of the afferent and efferent auditory pathways is necessary for proper auditory processing. The efferent system is in charge of both selective attention and the central control of cochlear amplification.

Speech and background noise could be difficult to separate in people with hearing loss, and selective attention is also impaired, which suggests that the efferent system is involved.

s-ABR may be used in the assessment of the auditory efferent pathways, which is crucial in difficult listening situations such as speech perception in noise and dichotic listening.

Being a simple and non-behavioral process, it can be used with children as a useful diagnostic tool for auditory processing disorders existing in a variety of disorders and provide differential diagnoses of diseases with similar symptoms. It can be used in different languages.

Acknowledgements

Not applicable.

Authors' contributions

HA analyzed the data and wrote the manuscript. NM reviewed the literature on speech-evoked ABR. DH edited the manuscript. EH made the final approval.

Funding

This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

Dr Ebtessam Nada is a co-author of this study and an Editorial Board member of the journal. She was not involved in handling this manuscript during the submission and review processes. The rest of the authors have no conflict of interest to declare.

Received: 4 April 2023 Accepted: 22 December 2023 Published online: 22 January 2024

References

- 1. Khuwaja A, Haghighi J, Hatzinakos D (2015): 40-Hz ASSR fusion classification system for observing sleep patterns. EURASIP J Bioinf Syst Biol 1:1–12
- 2. Jewett L, Williston S (1971) Auditory-evoked far fields averaged from the scalp of humans. Brain 94(4):681–696
- 3. Hall, J. (2015): eHandbook of auditory evoked responses, Kindle Edition. In Hall, M. (Ed.), Pearson Educatino Inc. Boston
- Picton W, Stapells R, Campbell B (1981) Auditory evoked potentials from the human cochlea and brainstem. J Otolaryngol 9(1):1–41
- Picton W (1992) The P300 wave of the human event-related potential. J Clin Neurophysiol 9(4):456–479
- Skoe E, Kraus N (2010) Auditory brainstem response to complex sounds: a tutorial. Ear Hear 31(3):302–304
- Johnson L, Nicol G, Kraus N (2005) Brainstem response to speech: a biological marker of auditory processing. Ear Hear 26(5):424–434
- Hornickel J, Knowles E, Kraus N (2012) Test-retest consistency of speechevoked auditory brainstem responses in typically-developing children. Hear Res 284(12):52–58
- BinKhamis, M. (2019): The speech auditory brainstem response as an objective outcome measure. (Unpublished thesis), The University of Manchester, United Kingdom.
- Anderson S, Skoe E, Chandrasekaran B, Zecker S, Kraus N (2010) Brainstem correlates of speech-in-noise perception in children. Hear Res 270(12):151–157
- 11. Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology 47(2):236–246
- 12. Hatzopoulos, S. (2017): Introductory chapter-genealogy of audiology. In Hatzopoulos, S. (Ed.), Advances in Clinical Audiology, . Intech Open. 3–28
- Moossavi A, Lotfi Y, Javanbakht M, Faghihzadeh S (2019) Speech-evoked auditory brainstem response: a review of stimulation and acquisition parameters. Auditory and Vestibular Research 28(2):75–86
- Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. Proc Natl Acad Sci 106(31):122–127
- 15. Johnson L, Nicol T, Zecker G, Kraus N (2008) Developmental plasticity in the human auditory brainstem. J Neurosci 28(15):400–407
- Kraus N, Nicol T (2005) Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends Neurosci 28(4):176–181
- Kaas H, Hackett A (1999) 'What' and 'where' processing in auditory cortex. Nat Neurosci 2(12):1045–1047
- Abrams, D. and Kraus, N. (2015): Auditory pathway representations of speech sound in humans. In Katz, J., Chasin, M., English, K. et al. (Eds.), Handbook of Clinical Audiology, Philadelphia: Wolters Kluwer Health. 527–544
- Liu F, Palmer R, Wallace N (2006) Phase-locked responses to pure tones in the inferior colliculus. J Neurophysiol 95(3):926–935

- 20. Parbery-Clark A, Tierney A, Strait L, Kraus N (2012) Musicians have finetuned neural distinction of speech syllables. Neuroscience 2(19):111–119
- 21. Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. Clin Neurophysiol 115(9):2021–2030
- Hayes A, Warrier M, Nicol G, Zecker G, Kraus N (2003) Neural plasticity following auditory training in children with learning problems. Clin Neurophysiol 114(4):673–684
- Killion C, Niquette A, Gudmundsen I, Revit J, Banerjee S (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America 116(4):395–405
- 24. Yamamuro K, Ota T, Iida J, Nakanishi Y, Matsuura H, Uratani M, Kishimoto T (2016) Event-related potentials reflect the efficacy of pharmaceutical treatments in children and adolescents with attention deficit/hyperactivity disorder. Psychiatry Res 2(42):288–294
- Kraus, S. and Canlon, B. (2012): Neuronal connectivity and interactions between the auditory and limbic systems. Effects of noise and tinnitus. Hear Res;288(12):34–46.
- Fujihira H, Shiraishi K (2015) Correlations between word intelligibility under reverberation and speech auditory brainstem responses in elderly listeners. Clinical Neuro-physiology 126(1):96–102
- Sanfins D, Hatzopoulos S, Donadon C, Diniz A, Borges R, Skarzynski H, Colella-Santos F (2018) An analysis of the parameters used in speech ABR assessment protocols. The journal of international advanced otology 14(1):100–102
- Song H, Nicol T, Kraus N (2011) Test–retest reliability of the speechevoked auditory brainstem response. Clin Neurophysiol 122(2):346–355
- 29. Song H, Skoe E, Banai K, Kraus N (2011) Perception of speech in noise: neural correlates. J Cogn Neurosci 23(9):268–279
- Anderson S, Parbery-Clark A, Yi G, Kraus N (2011) A neural basis of speechin-noise perception in older adults. Ear Hear 32(6):750–752
- Galhom D, Nada E, Ahmed H, Elnabtity N (2022) Evaluation of auditory efferent system using speech auditory brainstem response with contralateral noise. Egypt J Hosp Med 87(1):2064–2071
- Jenkins A, Fodor C, Preascco A, Anderson S (2018) Effect of amplification on neural phase locking, amplitude and latency to a speech syllable. Ear Hear 39(4):810–824
- Sanfins D, Colella-Santos F (2016) A review of the clinical applicability of speech-evoked auditory brainstem responses. Journal of Hearing Science 6(01):9–16
- Bogliotti C, Serniclaes W, Messaoud-Galusi S, Sprenger-Charolles L (2008) Discrimination of speech sounds by children with dyslexia: comparisons with chronological age and reading level controls. J Exp Child Psychol 101(2):137–155

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.