DOI: 10.61782/fa.2025.0720

FORUM ACUSTICUM EURONOISE 2025

# NEURAL RESPONSES AKIN TO ABRS MEASURED DURING NATURAL SPEECH LISTENING FOR HEARING AID ASSESSMENT

**Florine L. Bachmann**[1*] **Joshua P. Kulasingham**[2] **Kasper Eskelund**[3] **Martin Enqvist**[2]
**Jens Hjortkjær**[4,5] **Emina Alickovic**[1,2] **Hamish Innes-Brown**[1,4]
[1] Eriksholm Research Centre, Snekkersten, Denmark
[2] Automatic Control, Department of Electrical Engineering, Linköping University, Sweden
[3] Oticon A/S, Smørum, Denmark
[4] Hearing Systems, Department of Health Technology, Technical University of Denmark, Denmark
[5] Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital, Denmark

## ABSTRACT

This review outlines the current status and proposes future directions for using auditory brainstem responses (ABRs) to continuous speech as a tool in hearing aid assessment. In current clinical practice, ABRs measured via electroencephalography (EEG) in response to short stimuli, averaged across repetitions, are a cornerstone of objective hearing screening and evaluation. Recent advancements have demonstrated that ABRs can be estimated from EEG recorded during continuous speech listening by using linear models which map the auditory speech features, extracted using auditory models, to continuous EEG activity. This approach enables the simultaneous investigation of multiple auditory pathway stages using complex, ecologically valid sounds; advancing potential future hearing assessments that can be seamlessly integrated into daily life. Current findings indicate that responses similar to click-ABRs, with a salient wave V peak, can be estimated from natural continuous speech. These responses are sensitive to speech level, can be captured using insert earphones or in a sound-field environment, and can be obtained from aided older listeners with hearing impairment. While these advancements bring us closer to ecologically valid objective hearing device and fitting assessments, there are remaining hurdles to overcome, such as assessing ABRs to continuous natural speech during standard hearing device use.

**Keywords:** *Auditory brainstem response (ABR), speech, electroencephalography (EEG), temporal response function (TRF), hearing aid.*

## 1. INTRODUCTION

Estimating brainstem responses to natural speech could potentially provide essential information on how a given hearing device processing scheme fits individual peripheral hearing abilities, offering the potential to improve hearing aid algorithms and fitting. Such assessments would require a method that can quantify *early* responses to *continuous natural speech* in a *reliable* way, is *sensitive* to changes in the incoming speech stimulus, and can be applied in *hearing aid users* while they are *using standard hearing devices*.

Current objective assessments of early sound processing do not fulfill all these requirements. The auditory brainstem response (ABR) is often employed for capturing early neural responses to sounds. Traditionally, methods to estimate ABRs measure the electroencephalogram (EEG) while the listener is exposed to many repetitions of the same sound. The EEG responses to each sound are averaged across a reasonably high number of sound repetitions (typically thousands) to estimate a transient neural response. This limits the stimuli to short duration sounds such as clicks [e.g. 1], chirps [e.g. 2–4], tones [e.g. 5, 6],

*__Corresponding author__: flln@eriksholm.com.

or syllables [e.g. 7, 8]. The transient stimuli are poorly suited for assessing the speech-specific hearing device signal processing algorithms [9], and the isolated speech units fail to capture the complexity of natural continuous speech [10]. However, in the past decade, approaches to estimate brainstem responses to continuous natural speech have been developed, refined, and extended [11–17].

These methodological advances enable the estimation of ABRs to continuous natural sounds, capturing the response of the brainstem to natural speech (NS-ABR). Beyond its potential in hearing aid assessment, NS-ABR also holds promising applications in research and clinical diagnostics. NS-ABR offers the potential for simultaneous estimation of speech processing in higher stages of the auditory pathway, as well as pairing with behavioral measures of speech comprehension. Initial results show intriguing patterns of brainstem and cortical processing in the presence of hearing loss and amplification [13] that could deepen our understanding of speech processing along the auditory pathway. In clinical diagnostics, NS-ABRs could add ecological relevance by providing an objective measure of peripheral auditory encoding of *speech* rather than *click* or *chirp* stimuli, supporting the identification of potential barriers to speech perception. Furthermore, presenting familiar sounds may increase measurement comfort for listeners, potentially lowering agitation which could lead to reduced movement and muscle tension artifacts and thus improve data quality [12, 13].

Here, we focus on perspectives for using NS-ABR in hearing aid assessment, where it could help evaluate the effects of device software on brainstem speech processing under different real-life conditions, offering essential feedback on algorithms and personalized fits in varied listening environments. We review current knowledge about the NS-ABR in terms of optimized response computation, understanding the response, its reliability, sensitivity to sound changes, and investigate hurdles still to overcome towards its application in hearing aid assessment.

## 2. COMPUTING THE ABR TO CONTINOUS SPEECH: TRFS MEET AUDITORY MODELS

NS-ABRs are typically estimated using forward linear models that map broadband speech features onto the time-lagged EEG response, estimating a temporal response function (TRF), in analogy to previous cortical analyses of continuous speech responses [e.g. 18–20]. This TRF can be interpreted as the evoked response to a unit change in the speech feature, often yielding a prominent peak at latencies similar to wave V of the click-ABR [11, 12, 16]. As linear models are unable to capture nonlinearities in the auditory periphery, this approach has recently been refined by non-linearly transforming the speech waveform into a signal simulating neural activity at the brainstem level using auditory nerve models (ANMs) [14–16]. Employing ANMs before neural response estimation via the TRF has considerably improved model performance and response signal-to-noise ratio (SNR) [15]. These improvements have led to shorter acquisition times [15] and allow for response estimation not only to continuous speech presented via insert earphones, but also to music [14], or speech presented in the sound-field [16]. While adding an ANM increases response SNR, other optimizations of the computation method, or improvements in EEG acquisition SNR for these specific signals could further reduce measurement time.

Although continuous speech and music are often encountered in daily life, a considerable share of real-life sound environments also include mixed and overlapping signals, and pauses. In conversations for example, speech often overlaps, and longer silent pauses occur [21]. In addition, sounds with very different spectrotemporal characteristics than speech or music, such as for example traffic noise, contribute to shaping real-life soundscapes. Further research could shed light on whether the current NS-ABR approach needs extension for application throughout those different and complex sound environments.

## 3. RESPONSE ORIGIN, SENSITIVITY, AND RELIABILITY OF THE ABR TO CONTINUOUS SPEECH

As a foundation for NS-ABR applications, it is crucial to understand the origin, sensitivity, and reliability of the response. Response analogies can facilitate interpretation and knowledge transfer from the extensive body of research on traditional brainstem measures. The estimated TRF mapping continuous broadband speech features to neural signals has been interpreted as an ABR [11, 12]. Its prominent response peak has a latency similar to ABR wave V measured to clicks [11, 12]. This short latency of less than 15 ms suggest a brainstem origin [22], as transmitting the neural signal to higher stages along the auditory pathway requires more time. The prominent peak of the NS-ABR is also positively correlated with the click-ABR wave V in both amplitude and latency [11, 12, 16], with the inferior colliculi and auditory midbrain areas thought to be the main sources of click-ABR wave V

[23–25].

When computed using a suitable auditory model, NS-ABR showed sound level sensitivity similar to click-ABRs [17]. Both showed larger amplitudes and shorter latencies at higher sound levels, and these level-dependent latency changes were consistent across the speech and click measures on an individual level [17]. This further underlines the analogies between NS-ABR and click-ABR, and suggests level sensitivity of the NS-ABR crucial for potential applications. In analogy to effects previously observed for click-ABRs [26], first results with processed 'peaky' speech stimuli show effects of energetic masking [27]. The target speech stream was presented in an increasing number of competing speech streams, leading to decreased SNR with increasing competing streams. In tact with decreasing SNR, ABR peaks to continuous 'peaky' speech occurred at longer latencies and lower amplitudes [27] - potentially indicating a similar sensitivity to masking for NS-ABR.

Another important consideration for applications is the test-retest reliability. In older aided participants with hearing-impairment, initial results indicate high test-retest reliability for identical speech stimuli, but differences for different speech material produced by the same speaker, potentially suggesting a high dependency of the response on the frequency content of the used speech signals [12]. While more work is needed for a comprehensive overview of NS-ABR test-retest reliability and its sensitivity to different abilities and sound environments, first findings on both are promising.

## 4. TOWARDS APPLICATION IN HEARING AID ASSESSMENT

Beyond understanding the response origin, its sensitivity, and reliability, further research is needed to lay the groundwork for use of NS-ABR as a tool in hearing aid assessment. Crucially, there is a need to understand the NS-ABR in the target group for hearing aids: people with hearing impairment - including those using their devices at the time of measurement. As the prevalence of hearing impairment shows a steep increase with age, particularly above the age of 50 [28, 29], older adults are the primary target population that could benefit from this tool. Initial findings in older adults with hearing impairment indicate that clear NS-ABRs could not be identified when presenting speech at 65 dB without any amplification [13]. Given that ABRs to clicks presented at low sensation levels relative to individual's hearing impairment are drastically reduced in amplitude [30], this suggests that the chosen 65 dB speech presentation level was too low for the participants' hearing loss and the sensitivity of the applied simpler NS-ABR method, which did not yet incorporate an auditory model [13].

The next step towards applying NS-ABR in hearing aid assessment is investigating the response in hearing impaired listeners *with amplification*. The same work that presented unamplified speech discussed above [13] also presented amplified speech to the older listeners with hearing loss, applying a linear amplification algorithm to the speech and directly presenting it via insert earphones. Findings for NS-ABR are consistent with those for click-ABR wave V, showing longer latencies and reduced amplitudes in older individuals with hearing impairment receiving amplification compared to younger normal-hearing listeners [12, 13]. More research on the NS-ABR in listeners with hearing impairment across all age groups is needed for a comprehensive understanding of the response in this diverse population.

Although these first findings are promising, evaluating hearing aid performance in real-life soundscapes requires estimating NS-ABR in listeners with hearing impairment *while providing amplification through an actual hearing aid*. This necessitates *sound-field speech presentation*. As a first step, we demonstrated that measuring NS-ABR in the sound-field was feasible in normal-hearing adults [16]. Slightly longer EEG recording time was required due to room acoustics, but this effect was migitated by pairing the TRF approach with a powerful auditory model, improving response SNR and resulting in clear NS-ABRs in all participants [16]. The next step is to integrate these elements by quantifying NS-ABRs in older individuals with hearing impairment while providing amplification through a standard hearing aid receiving sound in the sound-field.

## 5. OUTLOOK

ABRs measured to continuous natural speech (NS-ABR) instead of clicks and chirps is an emerging method with promise for applications. Initial research indicates that NS-ABR is reliable [13], sensitive to both speech level [17] and likely masking [27], can be quantified in older listeners with hearing impairment when presenting amplified speech [13], and when presenting speech via headphones [11–14, 17] as well as in the sound-field [16]. These findings suggest its potential for hearing aid assessment.

However, further research is needed to assess the feasibility of NS-ABR in individuals with hearing impairment while using standard hearing aids. By providing objective feedback from neural brainstem auditory processing of natural speech, the NS-ABR could help assess hearing aid software, such as noise suppression algorithms, and their adaptation to the individual listening needs in auditory scenes close to real life. This could contribute to improved hearing aid designs and better outcomes for hearing aid users.

## 6. REFERENCES


[1] M. Don and J. Eggermont, "Analysis of the click-evoked brainstem potentials in man using high-pass noise masking," *The Journal of the Acoustical Society of America*, vol. 63, no. 4, pp. 1084–1092, 1978.

[2] T. Dau, O. Wegner, V. Mellert, and B. Kollmeier, "Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1530–1540, 2000.

[3] C. Elberling and M. Don, "A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses," *The Journal of the Acoustical Society of America*, vol. 128, no. 5, pp. 2955–2964, 2010.

[4] S. G. Kristensen and C. Elberling, "Auditory brainstem responses to level-specific chirps in normal-hearing adults," *Journal of the American Academy of Audiology*, vol. 23, no. 09, pp. 712–721, 2012.

[5] T. Suzuki, Y. Hirai, and K. Horiuchi, "Auditory brain stem responses to pure tone stimuli," *Scandinavian Audiology*, vol. 6, no. 1, pp. 51–56, 1977.

[6] D. R. Stapells and P. Oates, "Estimation of the pure-tone audiogram by the auditory brainstem response: a review," *Audiology and Neurotology*, vol. 2, no. 5, pp. 257–280, 1997.

[7] N. Russo, T. Nicol, G. Musacchia, and N. Kraus, "Brainstem responses to speech syllables," *Clinical Neurophysiology*, vol. 115, no. 9, pp. 2021–2030, 2004.

[8] S. Anderson, A. Parbery-Clark, H.-G. Yi, and N. Kraus, "A neural basis of speech-in-noise perception in older adults," *Ear and Hearing*, vol. 32, no. 6, pp. 750–757, 2011.

[9] S. Laugesen, J. E. Rieck, C. Elberling, T. Dau, and J. M. Harte, "On the cost of introducing speech-like properties to a stimulus for auditory steady-state response measurements," *Trends in Hearing*, vol. 22, p. 2331216518789302, 2018.

[10] L. S. Hamilton and A. G. Huth, "The revolution will not be controlled: Natural stimuli in speech neuroscience," *Language, Cognition and Neuroscience*, vol. 35, no. 5, pp. 573–582, 2020.

[11] R. K. Maddox and A. K. Lee, "Auditory brainstem responses to continuous natural speech in human listeners," *eNeuro*, vol. 5, no. 1, 2018.

[12] F. L. Bachmann, E. N. MacDonald, and J. Hjortkjær, "Neural measures of pitch processing in EEG responses to running speech," *Frontiers in Neuroscience*, vol. 15, p. 738408, 2021.

[13] F. L. Bachmann, *Subcortical electrophysiological measures of running speech [PhD Dissertation at the Technical University of Denmark]*. DTU Health Technology, 2022.

[14] T. Shan, M. S. Cappelloni, and R. K. Maddox, "Subcortical responses to music and speech are alike while cortical responses diverge," *Scientific Reports*, vol. 14, no. 789, 2024.

[15] J. P. Kulasingham, F. L. Bachmann, K. Eskelund, M. Enqvist, H. Innes-Brown, and E. Alickovic, "Predictors for estimating subcortical EEG responses to continuous speech," *Plos One*, vol. 19, no. 2, p. e0297826, 2024.

[16] F. L. Bachmann, J. P. Kulasingham, K. Eskelund, M. Enqvist, E. Alickovic, and H. Innes-Brown, "Extending subcortical EEG responses to continuous speech to the sound-field," *Trends in Hearing*, vol. 28, p. 23312165241246596, 2024.

[17] J. P. Kulasingham, H. Innes-Brown, M. Enqvist, and E. Alickovic, "Level-dependent subcortical electroencephalography responses to continuous speech," *eNeuro*, vol. 11, no. 8, 2024.

[18] E. Alickovic, T. Lunner, F. Gustafsson, and L. Ljung, "A tutorial on auditory attention identification methods," *Frontiers in Neuroscience*, vol. 13, p. 153, 2019.

[19] E. C. Lalor and J. J. Foxe, "Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution," *European Journal of Neuroscience*, vol. 31, no. 1, pp. 189–193, 2010.

[20] S. Geirnaert, S. Vandecappelle, E. Alickovic, A. De Cheveigne, E. Lalor, B. T. Meyer, S. Miran, T. Francart, and A. Bertrand, "Electroencephalography-based auditory attention decoding: Toward neurosteered hearing devices," *IEEE Signal Processing Magazine*, vol. 38, no. 4, pp. 89–102, 2021.

[21] A. J. M. Sørensen, T. Lunner, and E. N. MacDonald, "Conversational dynamics in task dialogue between interlocutors with and without hearing impairment," *Trends in Hearing*, vol. 28, p. 23312165241296073, 2024.

[22] T. W. Picton, D. R. Stapells, and K. B. Campbell, "Auditory evoked potentials from the human cochlea and brainstem.," *The Journal of Otolaryngology. Supplement*, vol. 9, pp. 1–41, 1981.

[23] A. Starr, "Correlation between confirmed sites of neurological lesions and abnormalities of far-field auditory brainstem responses," *Electroencephalography and Clinical Neurophysiology*, vol. 41, no. 6, pp. 595–608, 1976.

[24] A. R. Møller and P. J. Jannetta, "Interpretation of brainstem auditory evoked potentials: Results from intracranial recordings in humans," *Scandinavian Audiology*, vol. 12, no. 2, pp. 125–133, 1983.

[25] J. K. Moore, "The human auditory brain stem as a generator of auditory evoked potentials," *Hearing Research*, vol. 29, no. 1, pp. 33–43, 1987.

[26] G. Mehraei, A. E. Hickox, H. M. Bharadwaj, H. Goldberg, S. Verhulst, M. C. Liberman, and B. G. Shinn-Cunningham, "Auditory brainstem response latency in noise as a marker of cochlear synaptopathy," *Journal of Neuroscience*, vol. 36, no. 13, pp. 3755–3764, 2016.

[27] M. J. Polonenko and R. K. Maddox, "The effect of speech masking on the human subcortical response to continuous speech," *bioRxiv*, 2024.

[28] GBD 2019 Hearing Loss Collaborators, "Hearing loss prevalence and years lived with disability, 1990–2019: Findings from the Global Burden of Disease Study 2019," *The Lancet*, vol. 397, no. 10278, pp. 966–1009, 2021.

[29] World Health Organization, *World Report on Hearing*. 2021.

[30] S. Verhulst, A. Jagadeesh, M. Mauermann, and F. Ernst, "Individual differences in auditory brainstem response wave characteristics: Relations to different aspects of peripheral hearing loss," *Trends in Hearing*, vol. 20, p. 27837052, 2016.