# Summary of the N1-P2 Cortical Auditory Evoked Potential to Estimate the Auditory Threshold in Adults

**Guy Lightfoot, Ph.D.**[1]

## ABSTRACT

This article introduces the cortical auditory evoked potential (CAEP) and describes the use of the N1-P2 response complex as an objective predictor of hearing threshold in adults and older children. The generators of the CAEP are discussed together with issues of maturation, subject factors, and stimuli and recording parameters for use in the clinic. The basic methods for response identification are outlined and suggestions are made for determining the CAEP threshold. Clinical applications are introduced and the accuracy of the CAEP as an estimator of hearing threshold is given. Finally, a case study provides an example of the technique in the context of medicolegal assessment.

**KEYWORDS:** Cortical evoked response, hearing threshold, nonorganic hearing loss, N1-P2

**Learning Outcomes:** As a result of this activity, the participant will be able to (1) describe the clinical uses and limitations of the N1-P2 cortical auditory evoked potential and (2) list the main stimulus and recording parameters for the adult N1-P2 cortical auditory evoked potential when used in the clinic to estimate the hearing threshold.

Introduced clinically in the late 1960s and early 1970s as an objective predictor of hearing threshold in adults and older children, the N1-P2 response provides a historical perspective to cortical auditory evoked potentials (CAEPs). The N1-P2 CAEP remains a valuable clinical tool for adults and older children who are unable to participate or reliably respond in conventional/standard audiometric procedures.[1] Obvious client groups include adults pursuing medicolegal compensation for alleged occupational hearing loss,[2–5] patients

(frequently adolescents) with suspected nonorganic hearing loss, and adults with marked learning difficulties or dementia. The advantages of the CAEP over the tone-evoked auditory brainstem response include better frequency specificity, inclusion of higher neurologic centers, less reliance on patient relaxation, and, especially if specialized equipment is used, somewhat better accuracy and shorter test time.[6]

The N1-P2 response is considered an obligatory exogenous evoked (or event-related) potential; that is, it does not involve cognitive processing although it is affected by arousal level. As other articles in this issue will show, CAEPs may be recorded in infants but the classic N1-P2 response does not fully mature in its ability to predict auditory sensitivity until subjects reach their late teens.[7]

## RESPONSE GENERATORS

Before discussing the generators of peaks N1 and P2, it may be helpful to introduce P1, the vertex-positive peak preceding N1. Simplistically, P1 is thought to arise from the primary auditory cortex, but P1 also may have contributions from the hippocampus, planum temporale, the lateral temporal cortex, and neocortical areas.[8,9] P1 has a latency of around 250 milliseconds in very young infants,[10] reducing with full maturity to typically 50 milliseconds. Peaks P1 and N2 are less prominent in the mature CAEP, which is dominated by N1 and P2, on which this article will focus.

The adult N1 is thought to have three temporally overlapping components.[11] The primary contribution to the stimulus onset N1 is frontocentral with a peak latency of 100 milliseconds and slightly greater activity at the scalp over the hemisphere contralateral to stimulation. It has bilateral and tangential generators in the auditory cortices on the superior temporal lobe. The second component of N1 is maximally recorded at midtemporal electrodes and is therefore called the T-complex. It consists of a positive peak around 100 milliseconds and a negative peak around 150 milliseconds, and it is thought to originate in the auditory association cortex of the superior temporal gyrus with a radial orientation. The T-complex is much larger and slightly earlier over the hemisphere contralateral to stimulation, compared with the ipsilateral hemisphere. The third component of the stimulus onset N1 has a variable latency around 100 milliseconds and may be associated with widespread transient arousal to increase sensory and motor responses to sound. Dipole source estimation indicates two distinct source activities underlying N1 with latencies of 100 and 145 milliseconds, both symmetrically localized in Heschl's gyrus of both hemispheres. P2 also appears to have multiple generators, in the primary and secondary auditory cortex in Heschl's gyrus of both hemispheres.

The tonotopic organization and time course of maturation for N1 and P2 differ.[12] Whereas N1 continues to mature into the teens, P2 (initially fused with P1 in infancy, when N1 may not be evident) has near-adult latency by age 2 years.

## SUBJECT FACTORS

In addition to the issue of maturation of the response (see above, together with other articles in this issue dealing with infant testing), the level of subject arousal, alertness, and attention are known to influence the latency, amplitude, or variability of the N1-P2 response. In general, a larger or less variable response is recorded when the subject is asked to read a book or attend to the stimuli (e.g., count the stimuli) rather than simply sit with eyes open.[13,14] It is generally held that the N1-P2 response is attenuated by sleep (which is therefore to be avoided when testing), but many studies have shown a more complex picture, in which response latencies increase during the stages of sleep but the amplitude of the response may either increase or decrease, depending on the sleep stage.[15,16] Unless the stage of sleep is assessed and controlled, it is prudent to employ a test protocol in which sleep is avoided.

## STIMULUS

The CAEP may be evoked by the onset (from silence) of an audible stimulus, to stimulus offset,[17] or to an abrupt change (e.g., frequency, phase, temporal or level change) in an otherwise continual stimulus, where the response is

referred to as the "acoustic change complex."[18] The stimulus onset response, which is used to estimate audiometric sensitivity, is triggered by the onset of the stimulus, and any stimulus more than approximately 30 milliseconds after stimulus onset does not contribute to the response.[19–21] This has important implications for the choice of tone burst duration and for the use of speech tokens as stimuli. In brief, it is the level and spectral content of the first few milliseconds of the stimulus that matters, providing that this evokes a response. Tonal stimuli must have their onset and offset controlled with specified rise and fall times if broad-spectrum clicks are to be avoided. The optimum value for tone burst rise time is a compromise between the conflicting demands of frequency specificity (best for long rise times) and neural synchrony (best for shorter rise times).[22] In clinical practice, tone burst rise and fall times in the range 10 to 20 milliseconds are popular.

In a recent study, Bardy et al demonstrated that a stimulus with greater spectral complexity evokes an on average 31% larger CAEP in terms of amplitude.[23] They used bursts of tones comprising four simultaneous unrelated frequencies within an octave range. They argued that this is not simply a bandwidth effect; band-limited noise stimuli of the same bandwidth evoked a smaller CAEP, as did more conventional single-frequency tone burst stimuli. This stimulus holds the promise of reduced test times or greater precision of the N1-P2 threshold estimate.

## RECORDING PARAMETERS

The parameters employed in research and clinical practice often differ, reflecting their conflicting demands. In the former, investigating sometimes subtle effects, it is common to see the use of multiple recording channels, large numbers of sweeps, and wide filter bandwidths. In the latter, where clinical efficiency is important for response identification rather than response characterization, the recording parameters are chosen to maximize the speed of response-to-noise ratio improvement. Furthermore, the use of CAEPs to predict the auditory threshold also involves the determination of the stimulus level at which a CAEP is absent. What

follows is a discussion of the parameters used in clinical practice.

A single recording channel is adequate, with the noninverting electrode placed at or slightly anterior to the vertex and the inverting electrode at the mastoid. Either mastoid is acceptable, although a slightly larger response is often recorded using the contralateral mastoid.[24]

At near-threshold stimulus levels, the N1-P2 response has a spectral peak in the 2 to 5 Hz range (the reader is encouraged to consider the response as a single cycle of a sine wave and calculate its frequency from typical latencies of N1 and P2). A high-pass filter at 1 Hz and a low-pass filter at 15 Hz are effective in attenuating noise beyond the primary spectral range of the response. Amplifier gain is chosen so that an artifact rejection level of $\pm 50\ \mu V$ or $\pm 75\ \mu V$ is obtained. Unlike research studies in which a separate eye blink recording channel is used to eliminate artifacts, in clinical work, no eye blink channel is used. The recording time base (also referred to as recording epoch or sweep time) may be 500 to 1000 milliseconds. When values over 500 milliseconds are used, it is helpful to include a period of prestimulus recording, where no response will be seen. This facilitates an estimation of the residual noise in the recording and allows any candidate response to be seen as distinct from ongoing background activity.

The stimulus repetition rate for the averaging process is an interesting compromise between conflicting considerations. The maximum rate that can be used without suffering a decrement in response size is typically one stimulus every 10 seconds,[25,26] but for clinical efficiency, we care more about the rate at which the signal-to-noise ratio is most improved; the faster the averaging process, the greater the signal-to-noise ratio improvement. In adults, the optimum rate appears to be one stimulus every 1 to 2 seconds (0.5 to 1 Hz) with somewhat slower rates being preferable in infants.[19,27,28] If we stimulate faster than one stimulus every 10 seconds, then an interesting effect is seen, which has relevance to the number of sweeps we choose to use in the average. The response to the first stimulus in an averaging run has been preceded by a period of silence so suffers no effect of habituation;

successive responses are smaller as habituation occurs. Most of the habituation occurs in the first three or four stimuli,[29] and it is not unusual to see a response in the unaveraged recorded brain activity for initial stimuli. There is some evidence that reducing the predictability of the stimuli increases response amplitude. Several tactics have been investigated and found to be effective, including varying the interstimulus interval,[30] side of presentation, or frequency of the stimulus.[31,32] However, some studies have failed to observe such effects.[6,25] It is possible that an effect is seen only when used sparingly, otherwise the variability of the stimulus loses its attention-grabbing novelty.

The number of required sweeps or stimulus presentations within an averaging run depends on the size of the response. For supra-threshold stimulus levels, a few tens of sweeps may be sufficient to identify a response with a high level of confidence. However, testers must resist the temptation to terminate averaging for very low numbers of sweeps (<10) because they think they can identify a response; the risk of mistaking noise as a response is unacceptably high. At levels close to threshold, where the response is small, 50 to 100 sweeps may be necessary to provide a sufficiently high response-to-noise ratio. Similarly, the requirement for response absence usually involves there being no recorded candidate response, with a level of residual noise in the waveform that is sufficiently low to provide confidence that a small response is not obscured by noise. At stimulus levels that will define the CAEP threshold, it is important to replicate the level. In the case of response presence, replication allows the consistency of the response to be assessed whereas in the case of response absence, the difference between the waveforms may be used as a measure of residual noise.

There is no agreed value for the aspect ratio used to display CAEP waveforms but guidance for clinical practice for CAEP testing is in preparation by the British Society of Audiology, which recommends 100 milliseconds = 5 µV.

## RESPONSE CHARACTERISTICS AND INTERPRETATION

In common with evoked potentials of all modalities, the amplitude and latency of the response change with the magnitude of the stimulus; larger amplitudes and shorter latencies are associated with high-level stimuli.[33,34] Amplitude reduces and latency increases as the stimulus level approaches the CAEP threshold. Knowledge of these input-output functions can aid waveform interpretation. Some specialist CAEP systems offer an objective assessment of response detection, given in the form of a *p* value, and have been validated in clinical populations.[35] Values of 0.05 or less are usually used, where this is the probability of no response being detected. For systems without objective scoring, the tester should apply consistent and predefined criteria for response detection and response absence. Note that there are three, not two, possible outcomes: response detected, response absent, and inconclusive.

The criteria for response detection should include:

1. The response should have an appropriate waveform morphology, amplitude, and latency.
2. The response should be repeatable, as judged by similarity between replicates.
3. The response morphology, amplitude, and latency should follow the expected trend of smaller amplitudes and longer latencies compared with responses obtained for a higher-level stimulus, when available.
4. The response should have a sufficiently high response-to-noise ratio to provide the tester with a high degree of confidence that the response is genuine.

The response-to-noise ratio requires an estimation of the magnitude of both the response and the residual noise in the averaged waveform. The response may be taken as the N1-P2 peak-to-peak amplitude. The residual noise in an averaged waveform is provided by some systems. Where this is unavailable, one practical estimate is to superimpose a pair of replicated waveforms and subjectively judge the average gap between them, across the entire recording window.

The criteria for response absence should include:

1. There should be no likely response present; a possible response with a response-to-noise

ratio less than that needed for response presence is not sufficient to qualify for response absence.

2. The residual noise in the waveform(s) must be sufficiently low to be confident that a small response is not obscured by noise. A value of 2 μV has been suggested by Van Dun et al.[23]

It is not sufficient to say, "I can't see a response"; the tester must have a high degree of confidence that a response is genuinely absent. Waveforms that do not meet the above criteria for response detection or response absence must be regarded as inconclusive and take no part in the definition of the CAEP threshold. Resolving inconclusive levels normally requires further averaging but occasionally, small or odd-looking responses remain inconclusive even after further averaging.

## DEFINING THE CAEP THRESHOLD

The CAEP threshold is defined as the lowest level at which a response is detected; ideally, a response is absent at a level of 10 dB or less below this level and, ideally, a demonstration of a response 5 or 10 dB above this level. It is sometimes sufficient to obtain responses down to a certain level without the need to obtain a formal threshold, for example in clinical cases if responses are recorded down to 20 dB hearing loss (HL). Such results should be described using the format ≤20 dB HL. Conversely, if no response was recorded at any stimulus level up to, say, 100 dB HL, where response absence was demonstrated, the results should be described using the format >100 dB HL.

## ACCURACY AND LIMITATIONS

There is an average difference between CAEP thresholds and the pure tone audiometry (PTA) thresholds in cooperative subjects; this is known as a bias and is typically 5 to 10 dB (e.g., 6.5 dB was reported by Lightfoot and Kennedy and 7.5 dB by Ross et al[6,34]). The CAEP threshold suggests a slightly more elevated threshold than the PTA. The value of this bias will depend to some extent on the methods used for stimulus calibration, response

acquisition and analysis, the presence of any loudness recruitment, and the presence of certain comorbidities. A smaller bias has been observed for greater degrees of hearing loss,[6] an effect attributed to recruitment. If the bias is to be subtracted from the CAEP thresholds, the bias must be determined locally. It is technically valid to subtract the bias when predicting the PTA threshold but this subtraction must be stated in any report. After subtracting any bias, there will be a spread of values in the CAEP–PTA difference in cooperative subjects. After accounting for their 6.5 dB bias, Lightfoot and Kennedy found that 94% of threshold estimate differences were ≤15 dB.[6] Such information may allow a confidence range to be associated with CAEP results.

For example, a CAEP threshold of 50 dB HL is obtained. The bias is rounded to 5 dB. Subtracting the bias gives the best estimate of the PTA as 45 dB HL; there is 95% confidence that the PTA lies in the range 30 to 60 dB HL (45 ± 15).

An Achilles heel of the CAEP is that in a small proportion of cases, the CAEP threshold can overestimate any hearing loss by more than 20 dB.[36,37] This can sometimes be attributed to subject fatigue or drug effects but is usually inexplicable, cannot be anticipated, and to date there is no effective means of mitigating the effect.

## CASE STUDY

A man in his mid-50s worked for 11 years as a welder in a noisy shipyard. Industrial relations between management and the workforce were poor, and claims for compensation relating to occupational hearing loss were common. His union arranged for all similar members to be assessed and supported his claim for occupational hearing loss. Hearing protection was available but the man complained that wearing it made him unable to hear warning sirens. Pure tone audiometry performed on behalf of both the plaintiff and defendant showed a fairly flat pattern bilateral symmetrical sensorineural hearing loss ranging from 60 to 80 dB HL. Hearing aids had never been sought. Medical experts disagreed on causation; one claimed the hearing loss was a consequence

of noise exposure and the other suggested the flat audiometric pattern, if genuine, showed that it was unrelated. The possibility of nonorganic hearing loss was raised. An independent CAEP testing was arranged to provide an objective assessment of the degree and contour of the man's hearing loss. At the interview, there were signs of nonorganic behavior (exaggerated difficulty in following instructions and hypersensitivity during tympanometry and electrode attachment).

Fig. 1 shows the CAEP waveforms obtained with 1 kHz tone burst stimuli, suggesting CAEP thresholds of 15 dB HL on the right and 30 dB HL on the left. Using a bias figure of 5 dB, these results suggested behavioral thresholds of 10 dB HL on the right and 25 dB HL on the left. Tests at higher frequencies showed a modest high frequency loss, consistent with the combined effects of age and noise exposure.

The equipment used in Fig. 1 was developed by the author in collaboration with Cambridge Electronic Design Ltd. (Electronic Design Ltd., Cambridge, UK). The man's claim was upheld but his disability and associated compensation were based on the CAEP results.

## SUMMARY

The N1-P2 CAEP is a valuable but underused tool in the audiologist's armory. It is most useful in cases of adults and older children unable or unwilling to perform accurate pure tone audiometry. It is less affected by muscle activity and is more frequency-specific than the auditory brainstem response. Disadvantages include poorer precision of threshold estimation in infants and younger children and the lack of time-efficient software and objective CAEP detection in mainstream auditory evoked potential systems.
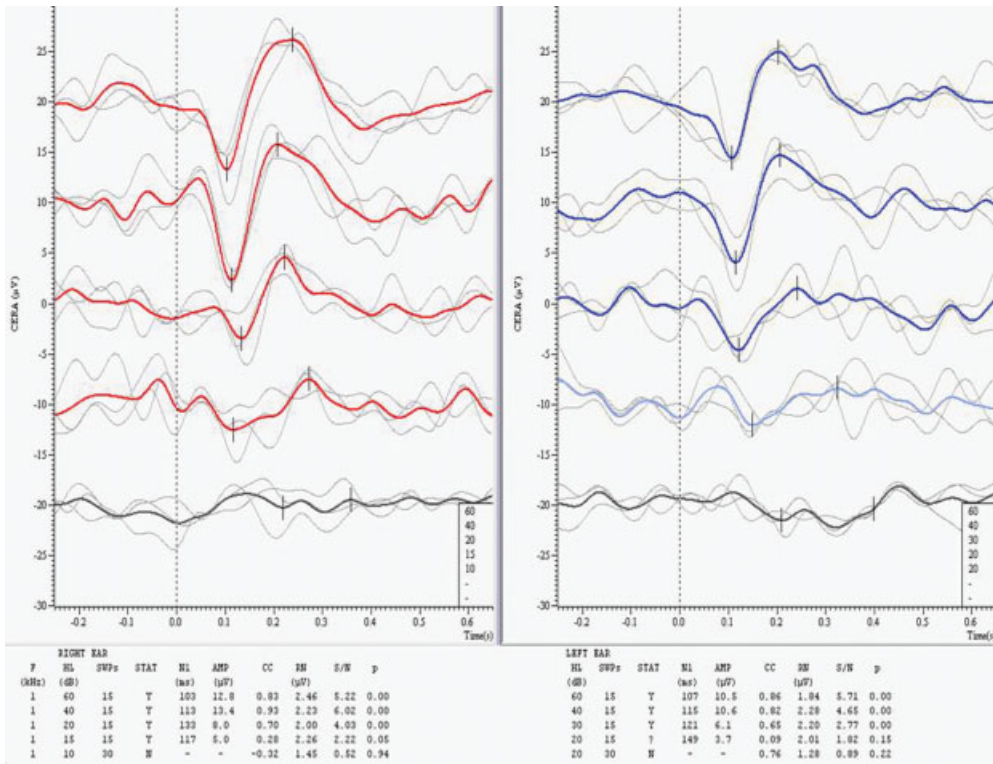
RIGHT EAR

| F (kHz) | HL (dB) | SWPs | STAT | N1 (ms) | AMP (μV) | CC | RN (μV) | S/N | p |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 60 | 15 | Y | 103 | 12.8 | 0.83 | 2.46 | 5.22 | 0.00 |
| 1 | 40 | 15 | Y | 113 | 13.4 | 0.93 | 2.23 | 6.02 | 0.00 |
| 1 | 20 | 15 | Y | 133 | 9.0 | 0.70 | 2.00 | 4.03 | 0.00 |
| 1 | 15 | 15 | Y | 117 | 5.0 | 0.28 | 2.26 | 2.22 | 0.05 |
| 1 | 10 | 30 | N | - | - | -0.32 | 1.45 | 0.52 | 0.94 |

LEFT EAR

| HL (dB) | SWPs | STAT | N1 (ms) | AMP (μV) | CC | RN (μV) | S/N | p |
|---|---|---|---|---|---|---|---|---|
| 60 | 15 | Y | 107 | 10.5 | 0.86 | 1.84 | 5.71 | 0.00 |
| 40 | 15 | Y | 115 | 10.6 | 0.82 | 2.28 | 4.65 | 0.00 |
| 30 | 15 | Y | 121 | 6.1 | 0.65 | 2.20 | 2.77 | 0.00 |
| 20 | 15 | ? | 149 | 3.7 | 0.09 | 2.01 | 1.82 | 0.15 |
| 20 | 30 | N | - | - | 0.76 | 1.28 | 0.89 | 0.22 |

**Figure 1** Cortical auditory evoked potential results at 1 kHz. Abbreviations: AMP, response amplitude; CC, cross-correlation coefficient; HL, hearing level; p, p value; RN, residual noise; S/N, response signal-to-noise ratio; STAT, status (Y=present; N=absent; ?=inconclusive); SWP, sweeps.

# REFERENCES

1. Hyde M. The N1 response and its applications. Audiol Neurootol 1997;2(5):281–307

2. Alberti PW, Hyde ML, Riko K. Exaggerated hearing loss in compensation claimants. J Otolaryngol 1987;16(6):362–366

3. Coles RR, Mason SM. The results of cortical electric response audiometry in medico-legal investigations. Br J Audiol 1984;18(2):71–78

4. Hone SW, Norman G, Keogh I, Kelly V. The use of cortical evoked response audiometry in the assessment of noise-induced hearing loss. Otolaryngol Head Neck Surg 2003;128(2):257–262

5. Prasher D, Mula M, Luxon L. Cortical evoked potential criteria in the objective assessment of auditory threshold: a comparison of noise induced hearing loss with Ménière's disease. J Laryngol Otol 1993;107(9):780–786

6. Lightfoot G, Kennedy V. Cortical electric response audiometry hearing threshold estimation: accuracy, speed, and the effects of stimulus presentation features. Ear Hear 2006;27(5):443–456

7. Stapells D. Cortical event-related potentials to auditory stimuli. In: Katz J, ed. Handbook of Clinical Audiology. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2002

8. Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr Clin Neurophysiol 1994;92(3):204–214

9. Howard MA, Volkov IO, Mirsky R, et al. Auditory cortex on the human posterior superior temporal gyrus. J Comp Neurol 2000;416(1):79–92

10. Sharma A, Dorman MF, Spahr AJ. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. Ear Hear 2002;23(6):532–539

11. Pratt H, Lightfoot G. Physiological mechanisms underlying MLRs and cortical EPs. In: Tremblay KL, Burkard RF, eds. Translational Perspectives in Auditory Neuroscience; Physiological Assessment of Audition. San Diego, CA: Plural Publishing; 2012

12. Wunderlich JL, Cone-Wesson BK, Shepherd R. Maturation of the cortical auditory evoked potential in infants and young children. Hear Res 2006;212(1–2):185–202

13. Mast T, Watson C. Attention and auditory evoked responses to low detectability signals. Percept Psychophys 1968;4:237–240

14. Keating LW, Ruhm HB. Within average variability of the acoustically evoked response. J Speech Hear Res 1971;14(1):179–188

15. Ornitz EM, Ritvo ER, Carr EM, Panman LM, Walter RD. The variability of the auditory averaged evoked response during sleep and dreaming in children and adults. Electroencephalogr Clin Neurophysiol 1967;22(6):514–524

16. Campbell KB, Colrain IM. Event-related potential measures of the inhibition of information processing: II. The sleep onset period. Int J Psychophysiol 2002;46(3):197–214

17. Pantev C, Eulitz C, Hampson S, Ross B, Roberts LE. The auditory evoked "off" response: sources and comparison with the "on" and the "sustained" responses. Ear Hear 1996;17(3):255–265

18. Jones SJ, Longe O, Vaz Pato M. Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of "streaming"? Electroencephalogr Clin Neurophysiol 1998;108(2):131–142

19. Cody DT, Klass DW. Cortical audiometry. Potential pitfalls in testing. Arch Otolaryngol 1968;88(4):396–406

20. Weber BA. Habituation and dishabituation of the averaged auditory evoked response. J Speech Hear Res 1970;13(2):387–394

21. McCandless GA, Best L. Summed evoked responses using pure-tone stimuli. J Speech Hear Res 1966;9(2):266–272

22. Skinner PH, Jones HC. Effects of signal duration and rise time on the auditory evoked potential. J Speech Hear Res 1968;11(2):301–306

23. Bardy F, Van Dun B, Dillon H. Bigger is better: Increasing cortical auditory response amplitude via stimulus spectral complexity. Ear Hear 2015;36(6):677–687

24. Ponton C, Eggermont JJ, Khosla D, Kwong B, Don M. Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. Clin Neurophysiol 2002;113(3):407–420

25. Nelson DA, Lassman FM, Hoel RL. The effects of variable-interval and fixed-interval signal presentation schedules on the auditory evoked response. J Speech Hear Res 1969;12(1):199–209

26. Davis H, Zerlin S. Acoustic relations of the human vertex potential. J Acoust Soc Am 1966;39(1):109–116

27. Rapin I. Practical considerations in using the evoked potential technique in audiometry. Acta Otol 1964;206:117–122

28. Appleby S. The slow vertex maximal sound evoked response in infants. Acta Otol 1964;206:146–152

29. Woods DL, Elmasian R. The habituation of event-related potentials to speech sounds and tones. Electroencephalogr Clin Neurophysiol 1986;65(6):447–459

30. Rothman HH, Davis H, Hay IS. Slow evoked cortical potentials and temporal features of

stimulation. Electroencephalogr Clin Neurophysiol 1970;29(3):225–232

31. Butler RA. The influence of spatial separation of sound sources on the auditory evoked response. Neuropsychologia 1972;10(2):219–225

32. Butler RA. Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. J Acoust Soc Am 1968;44(4):945–950

33. Picton TW, Goodman WS, Bryce DP. Amplitude of evoked responses to tones of high intensity. Acta Otolaryngol 1970;70(2):77–82

34. Ross B, Lütkenhöner B, Pantev C, Hoke M. Frequency-specific threshold determination with the CERAgram method: basic principle and retrospective evaluation of data. Audiol Neurootol 1999;4(1):12–27

35. Carter L, Golding M, Dillon H, Seymour J. The detection of infant cortical auditory evoked potentials (CAEPs) using statistical and visual detection techniques. J Am Acad Audiol 2010;21(5):347–356

36. Albera R, Canale G, Magnano M, et al. [Relations between pure-tone audiometry and cortical evoked auditory potentials]. Acta Otorhinolaryngol Ital 1991;11(6):551–562

37. Tsu B, Wong LL, Wong EC. Accuracy of cortical evoked response audiometry in the identification of non-organic hearing loss. Int J Audiol 2002;41(6):330–333