CHAPTER TWO

Physiologic Assessment of Young Infants – Puzzles and Challenges in Early Hearing Detection and Intervention

Martyn Hyde, PhD

Abstract

This paper touches on selected topics related to the validity, accuracy and efficiency of initial audiologic assessment of babies referred from newborn hearing screening. Utilization of OAEs, CMs, ABRs, ASSRs, and late cortical potentials are discussed. The puzzles and challenges discussed include the following: (1) Many factors can complicate the physiologic assessment and differential diagnosis of ANSD. (2) ABRs to rarefaction and condensation clicks can show extraordinary and puzzling differences, (3) slow cortical potentials can yield hearing threshold estimates in sleeping newborns, (4) the frequency specificity of the ASSR should not be based simply on stimulus energy spectra, (5) comparative studies of techniques should follow some simple methodological rules, and (6) ABR thresholds can be more efficient if stimulation and averaging tactics follow some basic statistical principles.

Accurate, Early ANSD Diagnosis – Sometimes Easy, Often Challenging or Impossible

ANSD is characterized by degraded neural function relative to cochlear receptor function. Diagnostic criteria have evolved from 'normal OAEs and absent ABRs' to 'present OAEs and/or cochlear microphonics (CM) and absent or abnormal ABRs'. Major factors that govern physiologic assessment procedure and diagnostic inference include:

- ANSD can involve inner hair cells (IHCs), their synapses and the cochlear nerve, in isolation or combined.
- Dysfunction components/subtypes range in severity from minimal to complete.
- It can occur concurrently with conventional cochlear (ccHL) and conductive hearing loss (CHL).
- Pathology can change over time, due to degeneration, recovery, neurodevelopment and neuroplasticity.
- ANSD-type selective IHC pathology can occur with variable severity at specific cochlear frequency regions, alone or combined with typical ccHL pathology elsewhere in the same cochlea (Amatuzzi et al., 2011)!

Given absent ABR, normal OAEs rule in ANSD and rule out both moderate or greater ccHL and CHL; mild ccHL is not ruled out. Present OAEs at isolated frequencies suggest ANSD but are not definitive. OAEs are usually abolished by slight CHL, with risk of missing ANSD. If tympanometry shows a peak, OAE absence suggests ccHL but with no peak, OAE absence is non-diagnostic; in either case, ANSD is not ruled out. In fact, the only finding that might ever rule out ANSD is a normal, complete ABR waveform for clicks at or below 90 dBnHL.

If OAEs are minimal or absent, CM is measured to rarefaction/condensation clicks and distinguished from artifact by insert tube clamping. Large (300+ nanovolts), oscillatory CM with absent ABR rules in ANSD, but small, brief CM does not. Moreover, significant CM does not rule out even severe ccHL: CM can arise from OHCs, from IHCs if OHCs are compromised, and from any cochlear region (Withnell, 2001). CM is notoriously variable across subjects, abolished by large CHL and

depends strongly on stimulus level. Low-pass cut-offs below 3 kHz can abolish high-frequency CM and large CM can be associated with various central lesions (Santarelli, et al., 2006.)

With ABR present, interpretation is even more difficult. Postsynaptic ANSD can yield a prominent early peak blend of CM asymmetry, true summating potentials and wave I, with abnormal or absent later waves. At maximal click levels, vestibular potentials may intrude. ABRs that look antiphasic for rarefaction/condensation clicks can arise due to cochlear excitation phase effects and be confused with prolonged CM.

More sophisticated protocols to disentangle receptor (CM/SP) and neural potentials are required. Options include using several click levels, high (91+/s) and low (~21/s) rates, increased bandwidth, CM quantification and addition/subtraction of averages. Category rating scales for ANSD likelihood and severity are also required.

Currently, the amplitude ratio of pure CM and ABR V-V' is a crude but useful indicator; when V is clear, CM/V ratios above 2 strongly suggest ANSD, but there is a likelihood continuum and morphologic factors also contribute. Comparative CM and ABR data in young infants are extremely limited (Shi et al., 2012) and are a research priority.

Extraordinary ABR Waveforms

As noted earlier, evaluation of rarefaction (R) and condensation (C) click ABR records is commonplace to explore cochlear microphonic (CM) and differentiate between sensory receptor and neural activity. It is usually assumed that whatever inverts when rarefaction and condensation click records are compared is CM, whereas what does not invert is either SP or neural. Conventional wisdom is that latency shift between (R) and (C) ABR tracings is due to preferential neuronal activation in the scala-media-negative CM phase. For tonepips, the time difference is a half-period of the carrier, ranging from 125 ms for a 4 kHz tonepip to 1 ms for a 500 Hz tonepip. Peak cancellation or smearing effects in alternated average are negligible at 4 kHz and 2 kHz and, near threshold, are irrelevant at 500 Hz.

Usually, ABR morphology is similar for the two tonepip polarities but in a small (unknown) proportion of cases, the differences are marked. ABR waves can differ in size or even be absent for one polarity and present for the other. Waves can also differ in latency by much more than a half-period of the carrier, even to the point of response cancellation in alternated averages. These phenomena appear to be quite common in adults with high-slope, high-frequency hearing losses; combined extra-tympanic electrocochleography and ABR are common clinical tools with extraordinary asymmetries of CM and ABR waves a frequent finding. Examples of major asymmetries of R and C click ABRs in young infants are shown in Figure 1.

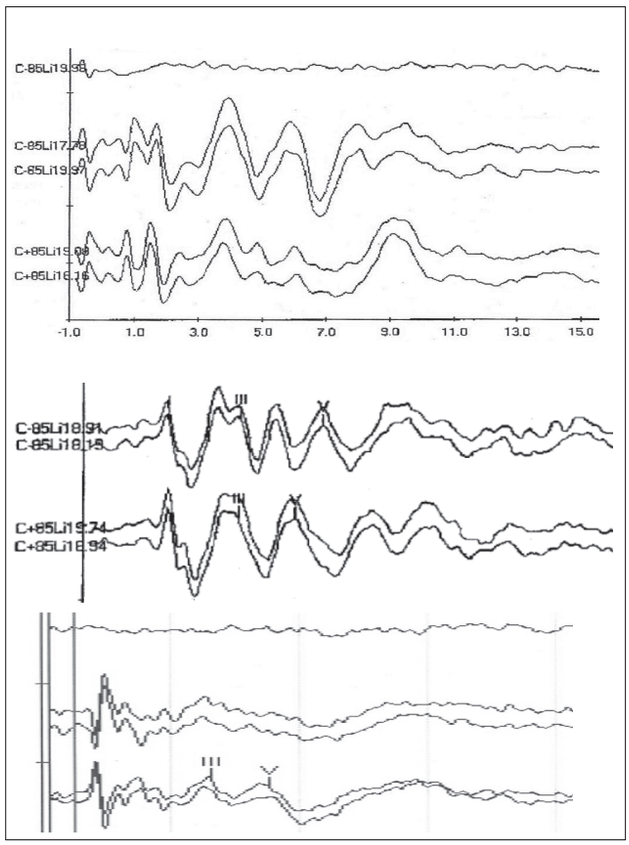

Figure 1. Clinical record extracts from three infants illustrating differences in rarefaction and condensation ABRs to 85 dBnHL clicks at 21.1 per second by insert earphone. X-scale 15 ms, Y-scale 0.5 V per division. All three blocks show replicated rarefaction responses with condensation responses immediately below; the upper and lower blocks also show tube-clamped averages. CM is apparent at about 1 ms in the upper and lower blocks.

The mechanism and pathophysiologic significance of these findings are unclear. A possible mechanism is differences in the (R) and (C) click travelling wave neural excitation profile and the resulting timing of ABR waves from different sites along the cochlear partition. Wave cancellation might occur if the place-sourced

component ABRs happened to be partially anti-phasic, which is more plausible than polarity-specific brainstem generator inactivity. It would be of interest to examine cases of ABR asymmetry using a subtractive, high-pass masking technique to isolate place-specific component ABRs. There may be place-ABR findings suggestive of cochlear dead zones.

Late Auditory Cortical Evoked Potentials – Forgotten Findings

Late (slow) cortical potentials, widely known in adult subjects as 'N1-P2', have undergone a resurgence of interest in the last decade. These late potentials have a number of possible applications such as biomarkers of cortical maturation, a measure of aided audibility and an objective index of perceptibility of a variety of acoustic events. These are clinical additions to the longstanding use of N1-P2 as arguably the best non-behavioural estimator of frequency-specific hearing thresholds in passively cooperative older children and adults. The N1-P2 has been especially useful in confirming functional hearing loss and in hearing assessments for medical-legal purposes (Hyde, 1997; Stapells, 2001). N1-P2 is not a unitary response; its two components differ in cortical site of origin, stimulus relationships and possibly their functional significance. The two features have different neurodevelopmental time courses, with N1 being absent in young infants.

N1-P2 has been described as an 'exogenous' response, reflecting its strong dependence on acoustic stimulus parameters such as rise rate, envelope, duration and energy-density spectrum (Hyde, 1997). Other late potentials such as the mismatch negativity and P300 are deemed to reflect more 'endogenous' aspects such as stimulus discriminability, probability and cognitive significance. However, N1-P2 is strongly dependent upon stimulus-oriented attention and in typical averaging sequences it is a highly habituated phenomenon dependent on stimulus repetition rate, mark-space ratio, regularity and sequence length.

In recent years, there have been numerous discussions about the 'acoustic change complex' (ACC). It would appear that the ACC is essentially the traditional N1-P2 response. Hearing threshold is actually a change from silence to sound, i.e., a special case of amplitude modulation. In fact, N1-P2 is evoked by any perceptible change in any dimension of the acoustic environment, including any direction of change in stimulus amplitude, frequency, spectral content and perceived source location.

The ACC terminology reflects obvious functional dynamics of N1-P2. Clynes (1969) reported on N1-P2 properties evoked by frequency-modulated (FM) stimuli. He found clear responses to upward and downward frequency changes in rectangular FM envelopes, but no responses at all to triangular FM, not even at the points of sharp change in direction of modulation. This led Clynes to describe N1-P2 functional dynamics in terms of a 'rest-motion (R-M) brain function'. That is, N1-P2 does not simply reflect temporally discrete acoustic change but only such change when it is preceded by a sufficient period of lack of change or 'sensory rest'.

In adults, N1 has typical latency about 100 ms and P2 about 180 ms, for tone bursts of moderate intensity and typical rise/fall and plateau times of at least 10-20 ms. Response amplitude follows an energy integration profile plateauing at about 30 ms. Repetition rates are typically 0.5 per second, with 20-50 stimuli per average. In infants, N1 is not well-developed and P2 is the main feature, typically peaking at about 300 ms but with substantial range.

In EHDI programs, P2 can be used to estimate hearing thresholds when the accuracy of ABR-based estimates is questionable, the obvious example being suspected or definite ANSD. Current clinical practise is to wait for ear-specific, frequency-specific visual reinforcement audiometry results (VRA), typically obtainable at age 6 months or later. However, this 6-month wait for results may be a major source of frustration and anxiety for families. Furthermore, waiting so long may incur significant, unnecessary delay or compromise in auditory neurodevelopment. For infants with visual, motor or cognitive comorbidities, reliable VRA results may never occur. If ANSD were suspected or identified at 1-3 months of age, the rationale for seeking an early alternative to behavioural threshold measurement is compelling. In major pediatric audiology centres, early use of P2 is increasingly common. Stimuli can be tone bursts by air or bone conduction, masked in the opposite ear when necessary and calibrated in dB HL.

There is widespread belief that P2 measurement in infants is best done with the infant awake. It is not clear how this viewpoint arose and whether it is valid. For estimating hearing thresholds in young infants, the strongest published report to date is based on testing in natural sleep. Taguchi, Picton, Orpin et al. (1969) reported 'excellent results' (almost 90% success) for tone burst threshold estimation at 500 Hz and 2 kHz in 250 newborns less than two weeks of age. Among the rationales for testing during sleep is that very young infants

sleep naturally most hours of the day and the auditory evoked responses are easier to see and measure with subjects asleep. Age is a major determinant of optimal P2 testing approach; the earlier the testing is done the more viable sleep testing becomes. Taguchi et al's results suggest that skill in recognising sleep state using ongoing EEG is relevant to determining accurate threshold estimation. It has long been known that the latency, size and morphology of AEPs depends on sleep state, so averaging across a mixture of conditional states may be misleading. The younger an infant is, the simpler it is to determine their sleep state categories by monitoring of the ongoing EEG morphology and the easier it is to render sleep state distinctions. It may be sensible to acquire relatively small averages along with their sleep state categories, then amalgamate the mini-averages in matching sleep states post hoc. Taguchi et al used tone bursts with 3-second cycle times; inter-stimulus recovery and sequence habituation time constants are larger in newborns and infants, so parameters that are optimal for adults will not apply.

The history of adult N1-P2 utilization and accuracy is typified by a wide range of technique, skills and understanding. Accurate work with late potentials is different tactically, much more interactive and generally more demanding than most ABR measurements. Ishida, Stapells & Small (in preparation, 2013) confirmed the ability to elicit the P2 response in sleeping infants in response to 60 dB white noise bursts in 16 sleeping young infants with presumed normal hearing (Figure 2).

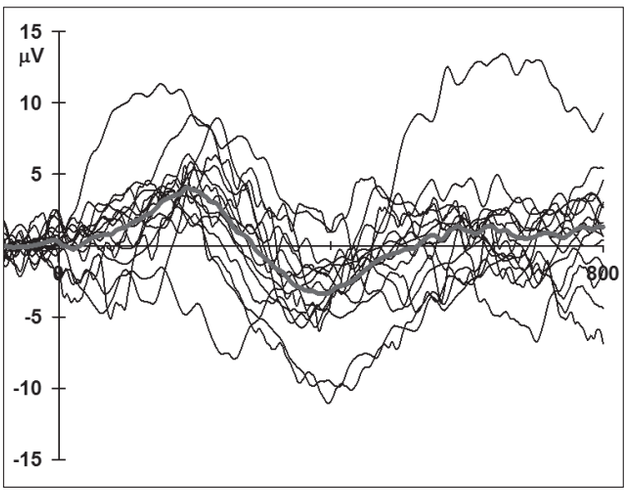

Figure 2. Slow cortical potentials elicited by 60 dB white noise bursts, from 16 sleeping infants. Waveforms from individual subjects plus the grand mean waveform are shown (Personal communication, figure extracted from Ishida, I., Stapells, D. & Small, S. (in preparation, 2013).

There are many other potential applications of N1-P2. For example, Michalewski et al. (2005) used N1-P2 in adults to measure gap detection thresholds. Dimitrijevic et al. (2012) found interesting differences in N1 amplitude to FM and AM stimuli at low and high tonal frequencies that interacted with ANSD pathologic subtypes. It remains to be seen what ultimately can be revealed in infants with this under-exploited tool.

Brainstem (80 Hz) ASSRs & ABRs

Another puzzling challenge relates to the ABR and the auditory steady state response (ASSR). Is the ASSR actually a tonepip ABR evoked by stimulus repetition at 80-110 Hz and analyzed in the frequency domain, and does it matter if it is or not? In this author's view, on current evidence the answer is 'sometimes' to both. At high sensation levels, energy periodic at the modulation frequency may include: stimulus artifact, cochlear receptor potentials, an entire ABR wave sequence from I through VI and rate-adapted remnants of later potentials such the middle-latency response (MLR). Depending on recording bandwidth, the slow postsynaptic potential components of the ABR also may be represented. Fourier analysis of response magnitude will not distinguish these sources, but they can usually be distinguished in the time domain.

At near-threshold intensity, the transient ABR evoked by a tonepip usually comprises a wave V peak and a trough often called V'. The peak-to-trough interval varies around 5 ms and there may be a wave VI with a V-VI interval of around 10 ms. For the ASSR, there is a sequence of brief quasi-sinusoids concatenated at the modulation frequency. Stimulus artifact, receptor potentials and MLR residues will vanish and slow postsynaptic components will be present or absent depending on bandwidth. Under these conditions, the phenomena may be assumed to be identical, unless and until contradictory evidence is found.

However, equality of generators does not mean equality of audiometric characteristics. Stimulus waveforms for the transient ABR and ASSR may be very different (Stapells, 2011). A common tonepip ABR stimulus is a 2-1-2 cycle trapezoidally modulated tone, and the ABR is an onset phenomenon governed by stimulus energy in the first 2-3 ms from onset. Above 2 kHz, the 2-1-2 cycle stimulus is too short to evoke optimum ABR amplitude, having energy-equivalent duration of only 0.58 ms at 4 kHz. In contrast, at 500 Hz only about the first half of the initial rise contributes to the ABR. For the simplest possible ASSR with sinusoidal modulation at 80 Hz, the

zero-to-peak time is 12.5 ms and only the first part of the rise generates the ABR waves giving rise to the ASSR. The effective stimulus magnitude is significantly less than the commonly measured levels, whether in peak SPL, RMS SPL or dBHL.

A related puzzle is the frequency-specificity of the tonepip ABR and ASSR. The energy density spectrum of typical ASSR stimuli is far narrower than that of the typical 2-1-2 cycle gated tonepip with the same carrier frequency. This is universally interpreted to mean that the ASSR is much more frequency-specific that the tonepip ABR. But, the logic of that step is difficult to grasp. First, if the ASSR is an ABR and the ABR is an onset phenomenon with an integration time of 2-3 ms, it doesn't matter what the ASSR stimulus looks like for each modulation after 2-3 ms from onset. Yet, the Fourier spectrum reflects the entire modulation period. No one would do a tonepip ABR with a 20 ms tonepip and expect it to be much more frequency-specific than a 2 ms tonepip – the flaw is obvious.

A second issue is that in the abnormal cochlea, frequency resolution is vastly inferior to that of the resonant peak when the cochlear amplifier is intact, reducing to the mechanical resolution of the basilar membrane itself, for cochlear losses greater than about 60 dB. Since what we hear is governed by whatever is transmitted through the cochlea to the central auditory system, the relevance of the acoustic Fourier spectrum of the stimulus to the actual frequency specificity of an evoked potential is difficult to comprehend. This does not mean that the ASSR is not more frequency specific, only that the usual argument for that belief is invalid.

Challenges with Comparative Studies

It is common to ask whether technique A or B is better for a given purpose such as estimation of thresholds in babies who have failed ABR hearing screening. There are several challenges with such a deceptively simple question:

1. What is meant by 'better'? The usual answer would be 'faster' or 'more accurate'. But, speed and accuracy have an inverse relationship, so either one must be fixed and the other evaluated or, far better, the trading relationship function between the two must be determined.

2. Accuracy itself has two parts: bias (mean error) and precision. Of the two, precision is vastly more important, because bias can be corrected by fixed adjustment. Precision has two major components: within-subject and between-subject variation. For this reason, at least two repetitions of every condition for every subject are a minimum.

3. A more difficult issue is optimization of each technique. Are the parameters of each procedure chosen to be representative of common practise, or to be credibly optimal? It is unlikely that common practise is also optimal. If representativeness is chosen, then the question asked is not whether A is truly better than B, but whether A as commonly practised is better than B as commonly practised. Because optimality is a function of so many parameters of subjects, procedures and contexts of practice, this question is by far the most practicable.

4. The literature is full of deficiencies of data analysis. A general remedy is to always tabulate in an appendix all pertinent data for all subjects, so that third parties can re-analyze as they see fit. A common issue in such an analysis is over-reliance on summary statistics. Statistical means are uninformative compared to distributions or scatterplots. Standard deviations are misleading unless data are normally distributed and contain no outliers. Correlation coefficients are the most opaque of all: they are strongly affected by data range and vastly different patterns of relationship can yield the same coefficient.

5. Even good reports may overgeneralize inferences based on statistical means, so wherever possible, results for individual subjects should be illustrated graphically. In general, it is the exceptions to average behaviour that that are far more informative than the average results themselves.

ABR Threshold Estimation Efficiency – So Much Wasted Time!

What is 'sufficient' audiometric information? One possible answer is threshold estimates at 2 kHz and 500 Hz within 10 dB, plus valid inference of hearing loss type. Whatever the answer, the session comes down to threshold estimates for various stimulus frequencies and routes. How can threshold estimation be optimized? The two main variables are (1) tactics of intensity selection and (2) signal averaging. Consider a single frequency-route for which no prior information is available, such as air conduction measures at 2 kHz.

Threshold estimation is a set of binary response detection trials. A crucial principle is that for any trial, the more equiprobable the outcomes, the larger the information gain. So what should be the starting intensity level? In a nil-risk baby, the positive predictive value of AABR screening failure is about 0.1, so a response at 30 dBnHL is much more probable than a non-response. We would have to go lower intensity levels to raise the non-response probability, but that would enmire us in minor middle-ear disorders, so testing should start at the minimum level of interest because any higher level is less informative. If we get clear responses at the initial intensity level, then changing frequency to 500 Hz or switching ears should be the next step in the ABR testing protocol.

However, if there is no response at 30 dB at 2 kHz, one choice is to change to bone-conduction testing at 2 kHz at 30 dB to quickly determine is this sensorineural or conductive hearing loss? The other option might be to increase the air conduction intensity level, to explore the severity of hearing loss. If we chose the latter option, the question becomes to what level? The answer depends on the a priori distribution of hearing thresholds in the population with hearing loss at 2 kHz. Given a range of about 30-100 dB HL and an assumption of uniform distribution, the test level for equiprobability is about 60 to 65 dB with a 5 dB correction factor for ABR threshold estimates at 2 kHz. Given no ABR response at 60 dB, the next level to test might be a jump to 80 dB. This strategy of large intensity ascending step size is very efficient. The least efficient strategy is to repeat ABR testing in 10 dB ascent (or descent) steps.

If there is no ABR response at 30 dB but a clear response at 60dB, should we next decrease intensity to 50 dB or to 40 dB? In fact, ABR trials are not binary, because we have information about response size and signal-to-noise ratios (SNR). Accordingly, a reasonable strategy is to decrease intensity to 40 dB if the ABR response at 60 was large or decrease intensity to 50 dB if the ABR response at 60 was small. These principles can be applied to 5 dB steps, but it is far more efficient and useful to get complete results in 10 dB steps before even considering 5 dB steps. In the author's experience, ABR threshold definition within 10 dB can be accomplished in infants with no response at 30 dB within three further levels or less the vast majority of the time.

There are five key points to consider in signal averaging tactics. (1) You cannot conclude that a response is "absent" if you could not have seen it even if it were present. Negative judgements require that residual noise levels in averages are below a criterion level. (2) Positive judgements are based on the SNR or, when the SNR is not overwhelming, the determination is made on tracing reproducibility. Thus, the final intensity threshold bracket upper levels should always be repeated. (3) Due to the root-n law of diminishing returns from averaging, fewer smaller signal averages are generally more useful than one large signal average. Signal averages should never be smaller than 500 sweeps because of the increased variability. (4) The total time spent on any given stimulus condition should be limited; if you cannot decide on response presence or absence within three averages totalling 6,000 sweeps, something is wrong; either you are chasing shadows or the EEG noise levels are unacceptably high. The best tactic in these circumstances is to increase intensity by 20-30 dB and attempt to identify a definite ABR tracing. (5) There are three categories of ABR response judgment: present, absent, and indeterminate. When you cannot decide with confidence that the response is present or absence, **do not** guess – additional measurement is obligatory.

References

Amatuzzi, M., Liberman, M.C., & Northrop, C. (2011). Selective inner hair cell loss in prematurity: a temporal bone study of infants from a neonatal intensive care unit. *Journal of the Association for Research in Otolaryngology, 12*(5), 595-604.

Clynes, M. (1969). The dynamics of evoked vertex potentials: the R-M brain function. In E. Donchin & D. B. Lindsley (Eds.), *Average Evoked Potentials. Methods, Results and Evaluations. NASA Special Publication SP-191* (pp. 363-374). U.S. Government Printing Office, Washington DC 20402.

Dimitrijevic, A., Starr, A., Bhatt, S., et al. (2011). Auditory cortical N100 in pre- and post-synaptic auditory neuropathy to frequency or intensity changes of continuous tones. *Clinical Neurophysiology, 122(3),* 594-604.

Hyde, M. (1997). The N1 response and its applications. *Audiology and Neurotology, 2*(5), 281-307

Ishida, I., Stapells, D., & Small, S. (personal communication, S. Small, 2013).

Michalewski, H., Starr, A,, Nguyen, T., Kong, Y., Zeng, F. (2005). Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. *Clinical Neurophysiology, 113(3),* 669-680.

Santarelli, R., Scimemi, P., Dal Monte, E., & Arslan, E. (2006). Cochlear microphonic potential recorded

by transtympanic electrocochleography in normally-hearing and hearing-impaired ears. *Acta Otorhinolaryngolica Italica, 26*(2), 78-95.

Shi, W., Ji, F., Lan, L., Liang, S., Ding, H., Wang, H., Li, N., et al. (2012). Characteristics of cochlear microphonics in infants and young children with auditory neuropathy. *Acta Oto-Laryngologica, 132*(2), 188-196.

Stapells, D. (2001). Cortical event-related potentials to auditory stimuli. In J. Katz, R. Burkard & L. Medwetsky (Eds.) *Handbook of Clinical Audiology, 5th Ed.* (pp. 378-406). Baltimore, MD: Lippincott, Williams & Wilkins

Stapells, D. (2011). Frequency-specific threshold assessment in young infants using the transient ABR and the brainstem ASSR. In R. Seewald & A. M. Tharpe (Eds.) *Comprehensive Handbook of Pediatric Audiology,* (pp. 409-448). San Diego, CA: Plural Publishing.

Taguchi, K., Picton, T., Orpin, J., & Goodman, W. (1969). Evoked Response Audiometry in newborn infants. *Acta Oto-Laryngologica, Suppl. 252,* 5-17.

Withnell, R.H. (2001) Brief report: the cochlear microphonic as an indication of outer hair cell function. *Ear & Hearing, 22,* 75-77.