Title: Cortical Auditory Evoked Potentials (CAEPs) in adults in response to filtered speech stimuli.

Authors: Lyndal Carter[1], Harvey Dillon[1], John Seymour[1], Mark Seeto[1], Bram Van Dun[1, 2]

[1] National Acoustic Laboratories, Sydney, Australia

[2] The Hearing CRC, Melbourne, Australia

Selected results from this paper were presented at the Audiology Australia XIX National Conference, Sydney, 16-19 May 2010.

Acknowledgements: The authors would like to acknowledge Maryanne Golding and Kirsty Gardner-Berry for their technical advice and support.

Submitted to: *Journal of the American Academy of Audiology*, 10th February 2012 (manuscript ID: 12-011)

Submitted as new manuscript:31 October 2012. Revised 18 December 2012.

ABSTRACT

Background: Previous studies have demonstrated that CAEPs can be reliably elicited in response to speech stimuli in listeners wearing hearing aids. It is unclear, however, how close to the aided behavioral threshold (i.e., at what behavioral sensation level) a sound must be before a cortical response can reliably be detected.

Purpose: The purpose of this study was to systematically examine the relationship between CAEP detection and the audibility of speech sounds (as measured behaviorally), when the listener is wearing a hearing aid fitted to prescriptive targets. A secondary aim was to investigate whether CAEP detection is affected by varying the frequency emphasis of stimuli, so as to simulate variations to the prescribed gain-frequency response of a hearing aid. The results have direct implications for evaluating hearing aid fittings in non-responsive adult clients, and indirect implications for evaluating hearing aid fittings in infants.

Research Design: Participants wore hearing aids while listening to speech sounds presented in a sound field. Aided thresholds were measured, and cortical responses evoked, under a range of stimulus conditions. The presence or absence of CAEPs was determined by an automated statistic.

Study Sample: Participants were adults (6 females and 4 males). Participants had sensorineural hearing loss ranging from mild to severe-profound in degree.

Data Collection and Analysis: Participants' own hearing aids were replaced with a test hearing aid, with linear processing, during assessments. Pure tone thresholds and hearing aid gain measurements were obtained, and a theoretical prediction of speech stimulus audibility for each participant (similar to those used for audibility predictions in infant hearing aid fittings) was calculated. Three speech stimuli, (/m/, /t/ and /g/) were presented aided (monaurally, non-test ear occluded), free field, under three conditions (+ 4 dB/ octave, **-** 4 dB / octave, and without filtering), at levels of 40, 50 and 60 dB SPL (measured for the unfiltered condition). Behavioral thresholds were obtained, and CAEP recordings were made using these stimuli. The interaction of hearing loss, presentation levels, and filtering conditions, resulted in a range of CAEP test behavioral sensation levels (SLs), from -25 to + 40 dB.

Results: Statistically significant CAEPs ($p < 0.05$) were obtained for virtually every presentation where the behavioral sensation level was greater than 10 dB, and for only 5% of occasions when the sensation level was negative. In these ("false positive") cases the greatest (negative) sensation level at which a CAEP was judged to be present was **-** 6 dB SL.

Conclusions: CAEPs are a sensitive tool for directly evaluating the audibility of speech sounds, at least for adult listeners. CAEP evaluation was found to be more accurate than audibility predictions, based on threshold and hearing aid response measures.

Key Words: Evoked potentials, auditory; hearing aids, adults.

Abbreviations: ABR= auditory brainstem response; ACA = aided cortical assessment; ANSD = auditory neuropathy spectrum disorder; ASSR= auditory steady-state response; ABS = Australian Bureau of Statistics; AIHW = Australian Institute of

Health and Welfare; BTE= behind-the-ear; CAEP = cortical auditory evoked potential; ; DSL= desired sensation level; 4FAHL = four-frequency average hearing level;; NAL = National Acoustic Laboratories; RECD = real ear to coupler difference; REIG = real ear insertion gain; SL= sensation level; SPL = sound pressure level.

Hearing aids provide a fundamental (albeit, partial) solution to the deficits associated with hearing loss (Dillon, 2001) by amplifying sound in order to make it audible to the listener. Restoring audibility is undoubtedly the clinician's most important goal in providing (re)habilitation to individuals with hearing impairment (Ching et al, 2001). Appropriate amplification is particularly crucial for infants with hearing impairment, as adequate reception of the speech signal is needed for the development of speech and language (Stelmachowicz et al, 2000). A number of early studies (described by Erber & Witt, 1977) demonstrated that children with moderate to severe hearing loss typically require the speech signal to be 20 to 40 dB above their individual speech detection levels, in order to attain maximum scores on auditory recognition tests.

Clearly, successful aural habilitation critically depends on providing audibility of the complete range of speech sounds (Ching, et al, 2001). However, audiologists working with pediatric clients will be well aware that, "…the problem comes in the implementation". "We can all agree that children need a safe, audible signal, but how do we fit a hearing aid to achieve this?" (Palmer, 2005).

Prescriptive formulas are widely used to determine an initial gain-frequency response for infants' hearing aids. These procedures prescribe gain on the basis of individual hearing thresholds (Dillon, 2001). This presents the primary challenge. When using electrophysiological methods to estimate hearing threshold level, there remains a level of uncertainty. Even once a child is old enough to cooperate in behavioral testing it may still be difficult to obtain precise and complete audiometric data (Stelmachowicz and Hoover, 2009). The relationship between hearing loss and required gain is based on certain assumptions, relied on in the derivation of the prescription formulae. Even when reliable thresholds can be obtained and hearing aids can be very precisely fitted to match prescriptive targets, there is still a need to evaluate the success of the fitting (Dillon, 2001). Unfortunately, "one can't ask infants the inevitable hearing aid question: *how does it sound?"* (Dillon, 2005).

The "difficult to test"

In addition to young infants, there are a significant number of older children and adults, sometimes referred to as the "difficult to test" (Ray, 2002), for whom it is problematic (or impossible) to accurately perform routine audiological assessments. The "difficult to test" population includes individuals who are variously referred to as intellectually impaired, intellectually disabled (Wen, 2007), or as "developmentally disabled". The "difficult to test" also includes individuals who do not experience functional limitations but are deliberately non-cooperative.

Australian census data (ABS, 1993a) indicated that around 1% of Australians (in all age groups), were reported as having an intellectual disability, and also requiring assistance with basic living activities, including verbal communication (Wen, 2007).

This is consistent with earlier US estimates (Ray, 2002). Additionally 1.4% of Australians were reported to have suffered brain injury or stroke (ABS, 1993b). It is suggested, based on 2003 data, that dementia affects between 6.5 and 7.4% of Australians over 65 (AIHW, 2007). There is also evidence that the number of affected adults is increasing. In 2006, almost 190,000 Australians were estimated to have dementia (Runge et al, 2009). According to the AIHW, the number of people in Australia affected by dementia will be 465,000 by 2031 (AIHW, 2007).

There is also evidence of an increase in the occurrence of disability in neonates, and while there have been remarkable improvements in birth survival rates, there has not been a concomitant decrease in long-term neurodevelopmental disability rates (Allen, 2008). A significant number of infants surviving preterm birth will experience long-term neurodevelopmental disabilities, including cerebral palsy and significant cognitive, visual, and hearing impairments (Allen, 2008). Studies have shown that 32% (Mace et al., 1991) and 27% (Fortnum et al., 2002) of hearing impaired children had at least one other handicap in addition to hearing impairment.

Hearing aid evaluation for the "difficult to test"

Traditionally, hearing aid evaluation for infants, or "difficult to test" clients has relied heavily on Behavioral Observation Audiometry (BOA). BOA is a subjective technique limited in its efficacy by the fact that responses are unlikely to occur consistently near "true" hearing threshold (Thompson and Weber, 1974). In contrast, CAEP assessment is an objective technique which does not rely on cooperation from the listener. CAEPs represent summed neural activity in the auditory cortex in response to sound. In adults, the CAEP consists of a positive peak (P1) around 50 ms followed by a negative deflection (N1) around 100 ms and another positive peak (P2) around 180 ms (Martin et al, 2007). Presence of the P1-N1-P2 complex indicates that the stimulus has been detected at the level of the auditory cortex (Hyde, 1997).

CAEP morphology is dependent on several factors. First, subject age, accumulated exposure to sound, and maturation of the auditory system, all determine the latency of the P1 positivity. The younger or less exposed to sound an individual is the longer its latency with respect to stimulus onset (Wunderlich, 2006). Second, sleep stage affects waveform morphology, with slower and larger recordable CAEPs if the subject is in deep sleep when compared with the awake state (Mendel, 1975). Third, attention increases amplitudes of N1 and P2 components (Picton, 1974). Fourth, longer stimulus lengths, higher intensities, lower frequencies, larger signal-to-noise ratios, and longer interstimulus intervals increase CAEP amplitudes and decrease latencies (Billings, 2009; Jacobson, 1992; Picton, 1977).

CAEPs can reliably be generated in response to speech stimuli for adults and children (e.g., Kurtzberg et al, 1988; Hyde, 1997; Purdy and Kelly, 2001; Cone-Wesson and Wunderlich, 2003; Purdy et al, 2005; Agung et al, 2006; Golding et al, 2006; Garinis and Cone-Wesson, 2007). There is also evidence that the CAEP response shows good agreement with behavioral thresholds for narrow-band stimuli (Rickards et al, 1996; Hyde, 1997; Tsui et al, 2002; Cone-Wesson and Wunderlich, 2003; Lightfoot and Kennedy, 2006). Of the electrophysiological techniques available, CAEPs have been regarded as most suited to assessing the audibility of hearing aid-amplified speech (Souza and Tremblay, 2006). CAEPs are generated at the highest level of the auditory pathway, and can provide physiological evidence that the speech signal has reached the cortex, and is thus potentially audible to the individual (Korczak et al, 2005). The response that arises from the auditory cortex is much larger (around 5 to 10 microvolts) than the amplitude of other electrophysiological measures (ABR or ASSR), so fewer stimulus presentations are needed for a result to be generated (Dillon, 2005). This presents an advantage in terms of required test time. CAEPs are also easier to record in a clinical setting, as larger waveforms are less susceptible to interference from other noise sources. However, Korczak et al (2005) noted that there are relatively few published reports of CAEP assessment of hearing aid wearers. Some studies have examined the effects of amplification on the CAEP, but solely in listeners with normal hearing (for example, Tremblay et al, 2006a, Billings et al, 2007; Billings et al, 2011). Few studies to date have systematically investigated the combined effects of sensorineural hearing loss and hearing aids on CAEPs.

There is, however, an extensive body of research pertaining to other applications of CAEP assessment, namely: for hearing threshold estimation, as an index of auditory system development (maturation), for suprathreshold assessments of auditory discrimination and speech perception, and to aid in the investigation of auditory neuropathy spectrum disorder (ANSD). In particular CAEPs can be used to measure the benefits of interventions (such as hearing aids and cochlear implantation). Sharma et al (2005) discussed three case studies in which children receiving hearing aids and/or cochlear implants, had CAEP assessment pre- and post-fitting. Sharma showed that the P1 latency can be used as an objective clinical tool to evaluate whether acoustic amplification for hearing-impaired children has provided sufficient stimulation for normal development of central auditory pathways. This study acoustically confirmed a previous study with children using cochlear implants

(Sharma et al, 2002), which showed that implanted children with the shortest period of auditory deprivation - approximately 3.5 years or less - evidenced age-appropriate latency responses within 6 months after the onset of electrical stimulation. These P1 latencies can also be used as a predictor of the implanted child's speech perception performance (Alvarenga et al, 2012) and the hearing aided child's auditory rehabilitation outcome (Thabet and Said, 2012).

Until recently, CAEP assessment in hearing aid evaluation has remained largely a research tool. This may, in part, be due to the lack of systematic data, and possibly an academic standpoint that too little is known about how the sound processing provided by amplification affects central auditory system encoding (Billings et al, 2007, Billings et al, 2011) and, subsequently, how it affects electrophysiological responses. Other factors may be the lack of ready access to practical test systems, and difficulties in interpreting the response waveforms. The shape of the CAEP response is variable, particularly in infants, and changes as the auditory cortex matures, through the teenage years and into early adulthood (Dillon, 2005). Carter et al (2010), however, demonstrated that a statistical technique (Hotelling's T^2) provides an efficient, automated method of response detection which is suitable for interpreting CAEP waveforms in infants.

Prior to the current study, there does not appear to have been a systematic study of audibility or the effects of more subtle variations in the hearing aid gain-frequency response. For this study, there were two main experimental hypotheses:

1. That, in adults with hearing loss, detection of cortical responses is consistent with aided behavioral sensation levels and predicted audibility calculations.

2. That cortical response detection may be affected when the frequency components of speech stimuli are altered using filters.

METHODS

Participants

There were 10 participants, 6 females and 4 males, all of whom had previously participated in a pilot test stage. Participants with various degrees of hearing loss (ranging from mild to severe-profound) and a range of audiometric configurations were selected. Pure tone thresholds for test ears are shown in Figure 1. All participants wore hearing aids regularly. Ages ranged from 39 to 82 years, with a median age of 67 years.

Informed consent was obtained. The experimental procedures were approved by the Australian Hearing Human Research Ethics Committee.

Test Device

To ensure consistency, a loan BTE hearing aid was substituted for the participant's own hearing aid(s) during the assessments. The device was a Siemens Triano S or, for participants with higher gain requirements, a Prisma 2 SP+. Participants wore their own earmold with the loan device. Earmold characteristics (style, material, venting etc.) varied among participants, according to the range of gain-frequency response requirements. Earmolds were not standardized, on the basis that CAEP detection is determined by the overall SPL in the ear canal provided by the hearing aid, not by the acoustic features of the hearing aid per se. The anti-feedback system and signal processing features (‘Hearing Comfort System’ or ‘Voice Activity Detection’) were disabled, and the compression scheme in all channels was set to linear in the loan device. While this is not usual clinical practice, this was deemed necessary in minimising unpredictable, automatic changes to the gain-frequency response of the device during assessments. The loan device was fitted as closely as possible to NAL-RP targets. Hearing aid measurements were performed using the Aurical system Real Ear Measurement (REM) and Hearing Instrument Test (HIT) modules. As illustrated in Figure 2, participants were reasonably well fit to the NAL-RP targets, within ± 5 dB of target at each measured frequency, except for three participants for whom the target gain could not be achieved at 3 and 4 kHz.

The maximum power output (MPO), as measured in the 2 cc coupler, was pre-set with consideration of the MPO of the participant’s own hearing aid. Assessments were monaural, in common with previous studies (Tremblay et al, 2006b; Billings et al, 2007; Billings et al, 2011), in order to simplify interpretation of the data. In all but two cases, the left and right ear hearing thresholds were symmetrical. In the cases of significant asymmetry, the better ear was the test ear. In cases of symmetrical hearing loss, the test ear was that participant’s preferred ear for monaural listening, or the right ear if there was no preference. The non-test ear was occluded using a foam hearing protector earplug throughout the protocol. Given that a poorer hearing ear was never used as the test ear, a more rigorous methodology for avoiding cross-heard signal (e.g., masking provided by insert phone to the non-test ear) was considered unnecessary.

Test Environment

Participants were seated in a high-backed armchair in a sound-proof booth. A loudspeaker positioned at 0° azimuth, 1.8 metres from the test position presented the test stimuli. An equalization filter corrected for the combined transmission response of the loudspeaker and room, as measured at the test point. A calibration check of the sound field was performed prior to each test session, using a Brüel & Kjær Type 2636 measuring amplifier and a ½ inch measuring microphone, suspended at the test position.

Stimuli

The stimuli used in this experiment were those available in the HEARLab® test system (ACA Module). They are recordings of the phonemes /m/, /g/, and /t/, which have been generated from natural speech tokens. The stimulus length for the /m/ and /t/ was 30 ms, and for the /g/ 21 ms. These phonemes were extracted from a recording of running speech, spoken by a female speaker. The stimuli were recorded with digitization rates of 40 kHz. They were gated off close to a zero crossing, to avoid audible clicks.

These essentially vowel-free stimuli were chosen because they have a spectral emphasis in the mid, low and high frequency regions respectively. These stimuli have been used extensively in cortical response projects at NAL (e.g., Golding et al, 2006; Golding, et al, 2007; Golding et al, 2009; Carter et al, 2010; Van Dun et al, 2012) for the assessment of adults and infants with normal hearing and those fitted with hearing aids. The stimuli were additionally filtered, according to the experimental design, using an Ultra Curve Pro, digital 24-bit converter. There were three filter conditions; a flat (unfiltered) response condition, and filter conditions of approximately + and - 4 dB per octave from 0 Hz to ~ 20 kHz. Each filter condition was applied to each of the three speech stimuli, resulting in a total of nine stimulus conditions.

Behavioral Assessments

Conventional pure tone audiometry was performed to confirm hearing threshold levels, using a 2-channel audiometer and TDH-39 headphones or 3A insert earphones, and a B71 bone conductor. The Hughson-Westlake procedure using a 5 dB step size was employed. Air conduction thresholds were obtained for each ear at the frequencies 250, 500, 1000, 2000, 3000, 4000, 6000 and 8000 Hz, with masking where appropriate. Bone conduction thresholds were obtained at 500, 1000, 2000 and 4000 Hz.

Behavioral aided thresholds for each of the speech stimuli (under all filter conditions) were measured in the same sound field as for CAEP assessment. The stimuli were presented via the HEARLab system, with a custom, continuous attenuator and external amplifier to allow the participant to adjust the signal level. Each stimulus was initially presented to the participant at an audible level. Then the attenuator was lowered by the tester so that the sound was below hearing threshold. The participant was then required to gradually increase the attenuator until the sound was just audible. Three ascending test runs were performed for each stimulus, unless the initial two runs were consistent to within ± 3 dB, in which case a third run was not performed.

The average of threshold values for each stimulus was taken as the behavioral threshold.

Audibility Calculation

In pediatric clinical practice, visual representations of the predicted amount of audible speech information are sometimes generated, e.g., the “speech-o-gram”, NAL-NL-1 fitting procedure, or “SPL-O-Gram” DSL fitting procedure (Scollie and Seewald, 2002, Frye and Martin, 2008), in order to describe the benefits (or limitations) of infant hearing aid fittings to parents and habilitationalists. These are calculated on the basis of the interaction between hearing thresholds, the gain-frequency response provided by the hearing aid, and the physical volume of the child’s ear canal, compared with an idealised long-term spectrum for average speech at a particular level (Dillon, Ching and Golding, 2008). The design of this study provided an opportunity to compare predicted (calculated) audibility (based on pure tone threshold levels) with behaviorally determined audibility (based on free field, aided hearing thresholds for speech stimuli).

Based on each participant’s pure tone audiometry thresholds, the predicted audibility of the three speech stimuli (at a 65 dB SPL presentation level) was calculated using the following procedure:

1. The spectral content of each of the three stimuli was measured in 1/3 octave band widths, for an overall level of 65 dB SPL (measured using an impulse time constant) in the free field. The levels in auditory bands centred at the same frequencies were calculated. For this calculation, auditory filters were

assumed to broaden with increasing hearing loss in the manner described by Moore and Glasberg (2004), under the assumption that 90% of hearing loss was due to outer hair cell loss, up to the specified maximum outer hair cell loss.

2. Individual ear, unaided pure-tone audiometry thresholds in dB HL for the frequencies 250, 500, 1000, 2000, 3000 4000, 6000, and 8000 Hz, were converted to the equivalent sound field level (in dB SPL) by adding the minimum audible field (MAF) for 0 dB HL (Bentler and Pavlovic 1989). These levels were increased by 4.8 dB and 6.1 dB, as an estimate of the effect of brief stimuli (30 ms durations, and 21 ms duration respectively) on hearing thresholds. This correction is based on an assumed energy integration time constant of 75 ms. Thresholds at all standard 1/3-octave frequencies from 125 to 8000 Hz were interpolated.

3. Measured hearing aid gains (Real Ear Insertion Gain, REIG) were subtracted from the unaided threshold levels at each frequency, to estimate the aided thresholds (dB SPL).

4. The resulting estimated aided thresholds were subtracted from the measured 1/3-octave band levels of each speech stimulus to derive an estimate of the sensation level at each frequency (calculated sensation level).

5. The calculated sensation levels across the different frequencies were compared to determine which 1/3-octave bandwidth had the highest sensation level, and this maximum figure was used as the final calculated (estimated) aided sensation level.

CAEP Assessment

CAEPs were recorded using a prototype of the HEARLab test system. As the standard HEARLab ACA module provides levels of 55, 65 and 75 dB SPL only, a custom attenuator and external amplifier were applied. The standard speech stimuli of the HEARLab ACA module were generated and presented from a free field speaker. The stimuli were alternated in polarity, and had an inter-stimulus interval of 1125 ms.

There were three stimulus presentation levels; 40, 50 and 60 dB SPL, the SPL being measured for each stimulus in the unfiltered condition. These presentation levels were chosen in order to produce a range of positive and negative sensation levels. These sensation levels were determined on the basis of the participant's free field behavioral threshold level for each test stimulus. These three levels resulted in nominally 270 test runs, comprising 3 speech sounds x 3 levels x 3 filter conditions x 10 participants, of which 45 were at negative sensation levels, 69 at sensation levels between 0 and 10 dB, and 147 at sensation levels greater than 10 dB.

Table 1 shows the breakdown of test runs by stimulus and behavioral sensation level.

Note that testing at the 40 dB SPL presentation level was not completed for one participant, as all 40 and 50 dB SPL presentations were at a negative sensation level (the 40 dB SPL presentations being markedly below threshold). A total of 261 recordings were therefore available for analysis.

The attenuator setting that produced the target SPL level in the unfiltered position was used for both the high boost (+ 4 dB/octave) and low boost (-4 dB/octave) conditions.

During the calibration procedure, the SPL level for each stimulus under the filtered conditions was recorded for later data analysis. The order of stimulus presentation was balanced among participants, according to a pre-determined protocol. There were no non-stimulus trials. As stated, however, the interaction of hearing threshold levels and stimulus presentation levels meant that test runs at below threshold levels occurred in 45 out of the total 261 runs (approximately 20%).

Participants were awake, and instructed to remain so, during testing. To help maintain a consistent state of alertness, participants watched a captioned DVD of their own choice throughout the testing. The tester observed the participant's state, via a video monitor, to ensure that they remained comfortable and suitably alert throughout the recordings. Participants were given a break if they appeared to lose alertness during the assessment.

Electrode sites were prepared by light abrasion with preparation gel, and were positioned as follows: active (non-inverting) at Cz (vertex), reference (inverting) at M1 (left mastoid), and ground at Fpz (forehead). Electrical impedance of the electrode contact was checked prior to recording and the electrode was re-applied if necessary to achieve an impedance of less than 5 kOhms. The target number of accepted epochs for each test run was 100 each for the /m/, /g/ and /t/ stimuli. However, because the HEARLab system presents the three stimuli in an interleaved paradigm, and stimuli may be rejected in different proportions, a slightly lower, or higher, number of accepted epochs (within 5 of the target) was occasionally allowed in the interests of limiting the total test time per participant.

Processing of Cortical Responses

During recording, the EEG activity was amplified in two stages. Firstly at the coupling to the scalp electrodes (x 121) and secondly (x 10) after the signal was transported through the electrode cables. The signal was down-sampled from 16 to 1 kHz and bandpass filtered online between 0.16 and 30 Hz. The recording window consisted of a 200 ms prestimulus baseline and a further 600 ms duration. Baseline correction was applied (calculated on 100 ms pre-stimulus). Because HEARLab is a clinical single channel system, it does not have the capability of eye blink rejection through an additional ocular channel. However, an artifact rejection criterion of about 100 μV has been adopted to reject all epochs that exceed a specific value, hence excessive noise sources (including eye movements) should be handled appropriately.

Response detection

EEG data were input to a MATLAB program to apply the same automatic detection algorithm as used in HEARLab for analysing adult CAEP responses. The waveform of each accepted epoch was averaged across successive 1-ms samples within each of nine time bins 33 ms wide. The first bin begins 51 ms after stimulus onset and the last ends 347 ms after stimulus onset. The resulting array of nine average values by 100 replicated epochs are then used as the input data for a Hotelling's T^2 (Flury and Riedwyl, 1988; Harris, 2001) statistical analysis. This analysis shows the probability of any linear sum of the value of the nine variables, averaged across epochs, being significantly different from zero.

Automated and machine scoring methods for detection of evoked responses are not new, but have not been routinely applied in the detection of cortical responses. The application of Hotelling's T^2 in cortical response detection has been recently reported using both adult- and infant-generated cortical responses (Golding et al, 2009; Carter et al, 2010). Results of these studies showed that Hotelling's T^2 was at least equal to, if not better than the average human observer at distinguishing genuine cortical responses from random electrical activity.

Although waveform morphology was not the main focus of this investigation, to calculate the N_1 and P_2 peak size and latency, the mean EEG was filtered with a 10Hz low pass filter. The maximum or minimum values, as appropriate, were selected from the regions of interest (50-150 ms post stimulus onset for N_1 and 150-250 ms post stimulus onset for P_2). Amplitude was measured as the average amplitude over regions of interest; 50 to 150 ms for $N_\mathbf{1}$, and 150 to 250 ms for $P_\mathbf{2}$.

RESULTS

Audibility / Sensation Level

Figure 3 shows the calculated sensation levels (predicted from the pure tone audiometric thresholds and gain frequency response of the test device, according to the audibility calculation procedure described previously) versus the behavioral sensation levels (as determined by obtaining the participant's hearing threshold levels for each speech stimulus in the free field). The correlation between the calculated audibility and the audibility measured behaviorally was high, at 0.84, suggesting that audibility calculation is a reasonable representation of likely audibility in the absence of objective aided measurements. In addition, absolute calculated and behavioral sensation levels are very similar, as demonstrated by the difference in calculated SL -

behavioral SL being approximately normally distributed with a mean of -1.1 dB (i.e., actual audibility was slightly greater than calculated), and standard deviation of 7.7 dB. This is illustrated in Figure 3, where it is evident that a line of best fit would intersect close to the origin of the graph.

CAEP Detection

Figure 4 shows the *p*-value (Hotelling's T^2 test of the null hypothesis of no cortical response) versus behavioral sensation level. With a detection criterion of $p < 0.05$, statistically significant CAEPs were detected for all but two test runs where the behavioral sensation level was greater than 10 dB. Additionally, there were only two cases in which the behavioral sensation level was negative and the *p*-value for detection of a CAEP was significant. It can be seen that the *p*-value systematically decreases as the presentation level increases above 0 dB SL. Figure 5 is the same plot for calculated sensation level.

Table 2 shows the percentage of CAEPs for which the detection *p*-value was less than 0.05 for all speech sounds in each of three behavioral sensation level ranges. A logistic regression model was fitted with CAEP detection as the dependent variable and sensation level, stimulus, and filter condition as predictor variables. It shows that although the likelihood of detecting a response is greatly affected by sensation level, it is not affected by the speech sound or filter condition (other than via the effect those factors have on behavioral sensation level).

Effect of Hearing Loss on Detectability

Figure 6 shows the *p*-value plotted against the four-frequency average hearing loss (4FAHL; average of 500, 1000, 2000 and 4000 Hz). Each vertical collection of points represents the data for different stimuli presented to a single participant. It is apparent that there is no systematic relationship between the detectability of responses and degree of hearing loss, once the behavioral sensation level of the stimulus is taken into account.

Cortical Response Amplitude

Figure 7 shows the grand mean of waveforms for all participants under all conditions, where the behavioral sensation level was greater than 10 dB.

Figure 8 shows a plot of P_2 – N_1 amplitude versus 4FAHL in the test ear. It is evident that there is no systematic relationship between 4FAHL and CAEP amplitude, once behavioral sensation level is taken into account. It does, however, appear that some participants inherently have larger CAEPs than others, even when sensation levels fall within the same broad range. Note that the amplitudes shown are those calculated by averaging the waveform within defined time intervals. These values are smaller than those that would be calculated by subtracting the true waveform peaks. This method of extracting peaks was used in order to reduce the biasing effects of residual noise in the waveform, which may be more problematic when the peak is defined as the maximum or minimum value occurring within a defined time period (refer to Figure 6.6 in Picton, 2011, for a more detailed explanation of this approach). Consequently, when it is not possible to reliably detect a cortical response, as has occurred in most instances for the participant in this study with profound loss, the measured response amplitudes cluster around 0 μV, and can include negative amplitudes.

A multiple regression model was fitted with $P_2 - N_1$ amplitude as the dependent variable and behavioral sensation level, stimulus, participant, and filter condition as predictor variables. Participant, stimulus and filter are categorical, while sensation level is continuous and is represented as a restricted cubic spline (Durrleman and Simon, 1989). The model used only those observations in which the stimulus was presented at a positive sensation level. The only factors found to have a significant effect on amplitude were participant, and the sensation level of the stimulus. To better determine how behavioral sensation level affects cortical response amplitude, a logistic model for the variation of amplitude with sensation level was fitted by non-linear regression. Based on theoretical expectations, the function's lower asymptote was fixed to be 0, and there was a separate upper asymptote parameter for each stimulus. All observations, including cases with a negative and a positive sensation level were included in the analysis. Figure 9 shows the data together with the fitted curves for each of the three test stimuli /m/, /g/ and /t/. A strong level effect was revealed, amplitude increasing as sensation level increases, even for sensation levels as low as 5 dB.

Cortical Response Latency

Only data points having a *p*-value less than 0.05 and N_1 or P_2 latencies not equal to the extreme values 52 ms and 150 ms (for N_1), or 152 ms and 250 ms (for P_2) were included in the analysis.

Sensation level

A regression line was fitted to N_1 and P_2 latency as a function of sensation level, without any adjustment for other variables. For brevity these figures have not been included. It was evident that both N_1 and P_2 latencies decreased as behavioral sensation level increased. N_1 latency shortened from 130 ms close to threshold to 100 ms at 40 dB SL, and similarly P_2 latency from 230 to 200 ms. A multiple regression model was fitted with N_1 and P_2 latencies as dependent variables, and sensation level, stimulus, participant, and filter condition as predictor variables. Behavioral sensation level, stimulus, and participant were all significantly associated with N_1 and P_2 latency. Filter condition was only found to be significantly associated with N_1 latency.

Table 3 shows latency difference estimates and their 95% confidence intervals for stimulus and filter. For stimulus, /m/ is the reference, and for filter, low-boost (LB) is the reference condition, e.g., /g/ is predicted to have a latency 11.3 ms less than /m/, and a 95% confidence interval for this difference is (-15.0, -7.6).

Residual Noise

Residual noise in the averaged waveforms was calculated on the basis of the epoch-to epoch variations in the recorded responses. The noise was estimated for each sample point in the average waveform for each participant and stimulus by calculating the variance, across epochs, around the average value at this sample point. This process was repeated across all sample points in the average waveform and the individual variance estimates were averaged, and then divided by the number of epochs. The square root was taken to give the resulting residual noise as an rms value in µV.

As there were minor variations in the actual number of accepted epochs recorded, an adjustment was made to account for this. This adjustment was achieved by multiplying the residual noise value by the square root of X, where X is the number of epochs divided by 100. A histogram of the resulting residual noise values is shown in Figure 10. The mean is 1.23 µV. The count indicates the number of test runs in which the residual noise was in each particular range (indicated on the X axis).

Figure 11 indicates the amount of residual noise as a function of 4FAHL. Residual noise does not vary systematically with 4FAHL, however there is clearly a significant inter-participant variation with some participants displaying noise levels 50% greater than other participants. When plotted, there was no suggestion of any association between cortical amplitude and residual noise amplitude.

DISCUSSION

Audibility and CAEP Detectability

Effect of sensation level

This study has demonstrated a close relationship between the behavioral sensation level of speech sounds and the presence of an aided cortical response, albeit in adult listeners wearing a linear device. The use of linear processing in this experimental design may cast doubt on the applicability of the findings to devices with more current signal processing. The presence or absence of the CAEP, however, is determined by audibility, and providing audibility is the goal of hearing aid fitting, regardless of the signal processing strategy employed. When the behavioral sensation level exceeded 10 dB, cortical responses were detected for 145 out of 147 stimulus presentations, corresponding to a sensitivity of 99%. In these two instances of CAEP non-detection (both in the high-boost condition), *p*-values did not reach the level of statistical significance; in the first case for /g/ at a behavioral sensation level of 10.1 dB, and in the second, for /t/ at a sensation level of 30.3 dB. In both cases, increasing the presentation level by 10 dB resulted in response detection, and in the latter case, a response was also detected at a 10 dB lower sensation level (20.3 dB). Although these individual cases are anomalous, taken in context this level of inconsistency would be tolerable in terms of making a useful clinical interpretation. Clinical implications are discussed in more detail later in this section.

As would be expected, for intermediate behavioral sensation levels (from 0 to 10 dB), the sensitivity averaged across this range is lower, (at 39/69 = 57%). Figure 4 illustrates that within this range of sensation levels, the actual sensitivity increases markedly as the sensation level increases from 0 to 10 dB.

Consistently, when the stimuli were presented below behavioral threshold for each stimulus, a "significant" cortical response (false positive response) was detected (with $p < 0.05$) for only two out of 45 stimulus presentations. This false detection rate of 4.4% is within confidence limits of 5% that one expects to occur when a detection criterion of 0.05 is used. In both false positive cases the *p*-values were only slightly below 0.05 (0.02 and 0.04 respectively). Although it is expected that 5% of detections will be invalid when a criterion of 0.05 is set, it is also possible that in these two cases the detections were not spurious, as the behavioral sensation levels were only slightly negative (-2 and -5 dB respectively), leaving open the possibility that cortical activity actually was elicited.

Scope of the findings

It is acknowledged that in this study there were a large number of stimulus presentations, but a small number of (adult) participants. This has three implications for the generalizability of the results. Firstly, the Hotelling's T^2 statistic used for response detection simply looks for a consistent response presence to each stimulus presentation. On this basis, the confirmation that the specificity is consistent with the adopted statistical criterion for response presence should be generalizable to any person tested. If no cortical response is present, the characteristics of the individual being tested should not affect the statistical process used to decide on response presence. Secondly, the extremely high sensitivity observed in this experiment is unlikely to generalize to people of all ages, nor even to all adults. It is evident from other studies that for some adult individuals, cortical responses cannot reliably be observed until the sounds have a much greater behavioral sensation level than is generally observed for most people (Hoth 1993; Tsui et al, 2002). In infants, Van Dun et al (2012) showed that in one out of four recordings (hence only a 75% sensitivity) a CAEP could not be detected when behavioral sensation levels were above 10 dB SL. Thirdly, when considering infants, as their waveforms are morphologically different from adult waveforms and change significantly with age and sound exposure (Sharma 2002), CAEP identification (aided versus unaided) and tracking of the effects of hearing aid use on the CAEP, might be more difficult.

Influence of other factors

In contrast to the marked effect of behavioral sensation level on CAEP detectability observed in this experiment, the other factors examined had little (or no) effect. While in principle, hearing loss, the type of speech sound, its SPL, and the way it is spectrally altered by filtering (which may represent the case in which a fitted hearing aid response is slightly less than optimal) will affect the likelihood of a cortical response being observed, it appears that these impact on response detection only via the effect these factors have on the behavioral sensation level of the sound, at least for the range of sounds used in this experiment.

Souza and Tremblay (2006) note that "Hearing aids modify the physical characteristics of sound; they introduce noise, compress signals, and alter the frequency content of the signal", and, "the effects of hearing aid processing on the physical characteristics of the sound likely affect the evoked neural response pattern". It is acknowledged that the hearing aids used in this study had the signal processing features disabled, and did not incorporate wide dynamic range compression. However, the findings indicate that with linear amplification, a frequency response close to prescriptive target, and a presentation level at least equivalent to soft- to-average conversational speech, that CAEPs should be expected - at least in cooperative adult listeners, with all but severe-profound hearing loss. It is also acknowledged that there were a number, albeit small, of false negative responses, even in this group of participants tested under ideal conditions. A limitation of the experimental design (due to time constraints) was that there was only one test run per stimulus/condition. It is unknown whether unexpected response outcomes (i.e., false negatives) would have re-occurred if the same test run was repeated. Because CAEPs are influenced by attention (Picton and Hillyard, 1974), it is conceivable that CAEP detectability may be affected by attention shifts, even within a test run. There is obviously value in repeat testing in cases of unexpected results, and also in obtaining results for a number of stimuli at a range of presentation levels, prior to making clinical interpretations. By obtaining *p*-values for different intensities, even where there is a degree of anomaly, one may be able to make an educated guess as to where the actual (aided) CAEP threshold might be (Picton and Hillyard, 1974).

CAEP Waveform Characteristics

As a statistical technique was used for response detection, the CAEP waveform characteristics were not directly relevant to the research aims of this study. However, some interesting observations were made in the course of the data analysis.

It is evident, both from Figure 8 and from the multiple linear regression analysis, that some individual participants have larger amplitude cortical responses than others. Not surprisingly, examination of Figures 6 and 8 suggests that those participants for whom cortical responses were most easily detected had the largest response amplitudes.

The variation of cortical response amplitude with behavioral sensation level as summarized by the logistical regression functions in Figure 9, is noteworthy in two respects. First, the growth in amplitude with sensation level commences as soon as sounds rise above the behavioral threshold. Second, the response reaches its maximum value by a sensation level of around 20 dB, a plateauing effect that has been confirmed in other studies (Picton, 1977; Ross, 1999). Finally, in regard to characteristics, it is also evident from Figure 11 that some participants consistently have higher residual noise levels than others, regardless of the fact these participants were ideally quiet and cooperative.

Audibility Calculation versus CAEP Measurement

It may be asked whether it is really worth measuring cortical responses in the very young or "difficult to test", in order to determine whether speech sounds are audible, assuming that this could adequately be predicted from knowledge of spectrum of the speech sound, the person's pure tone hearing thresholds, and the amplification characteristics of the hearing aid. Unfortunately, even for a simple linear hearing aid, as used in this experiment, prediction of audibility is not an accurate process. In addition to requiring a precise estimation of hearing threshold, prediction of aided audibility relies on a number of assumptions, including how the duration of brief sounds should be allowed for, the auditory filter bandwidth over which the stimulus intensity is integrated, the manner in which intensity falling above (or even below) threshold combine across different auditory filters, and the detectability of complex sounds relative to tonal sounds once these other factors have been considered.

Reliable behavioral thresholds were available for the participants in this study, however the inherent inaccuracy in predicting audibility is illustrated by the spread of data points in Figure 3. In principle, the discrepancies between calculated and behavioral sensation levels could arise from errors in either quantity. When the cortical detection *p*-values are plotted against calculated sensation level (Figure 5), a much less clear relationship between detectability and sensation level is evident. The relationship between cortical *p*-value and behavioral sensation level (Figure 4) is also much more orderly than the relationship between cortical *p*-value and calculated sensation level, strongly suggesting that the considerable spread in the relationship between calculated and behavioral sensation levels is caused by error in the calculated sensation level, rather than error in the behavioral sensation level. Using the cortical response as the gold standard, we thus infer that the spread of data points evident in

Figure 5 are much more likely to be due to errors in the calculation of sensation level than to its behavioral measurement.

It is important to note that the audibility calculations in this data are based on reliable pure tone behavioral thresholds from adult participants. Predictions will likely be less reliable when estimated threshold data is used.

Implications for Hearing Aid Evaluation

Despite growing interest in the use of the CAEP in the last decade, there has been continuing scientific debate about the appropriateness of the CAEP for hearing aid evaluation. The results of this study provide positive validation of the CAEP as an aided assessment tool, as at the level of average to slightly softer-than-average speech (50 and 60 dB SPL), CAEP responses for all three stimuli, under the unfiltered condition (which represented the NAL-RP response), were detected for all but two participants with reasonably well-fitted hearing aids. It is noteworthy that these two participants had the most significant hearing loss among the group, and consequently behavioral sensation levels were the closest to threshold among the test group.

It was interesting to observe in one of these two cases that the CAEP findings corresponded with the participant's hearing aid response preference. This participant consistently prefers gain well in excess of target (approximately + 15 dB REIG at 500 Hz) for everyday listening. For the purposes of the experiment the test hearing aid was fitted according to NAL-RP target, which resulted in less low frequency REIG than normally worn. Wearing the NAL-RP prescribed response, a CAEP for the /m/

stimulus in the unfiltered condition was not detected even at the highest presentation level, while a response *was* detected in the low-boost condition at 60 dB SPL. The CAEP reflected that an alternative response to the prescribed target provided better audibility for this individual.

It is also reassuring from a clinical point of view that, predictably, as the presentation level was decreased, the number of detected responses also decreased. At 40 dB (unfiltered), for the /m/, /g/ and /t/ stimuli respectively, only 3/9, 4/9 and 5/9 significant responses were observed. Responses at all three presentation levels (unfiltered) were detected for only two participants who, predictably, had the mildest degrees of hearing loss among the test group. From a clinical stand-point, this result also suggests that it is practical to set a lower limit to presentation levels for free field CAEP testing, as even for well fitted devices under optimal conditions, responses at < 50 dB SPL may only be observed in a small proportion of cases. Apart from this, the noise floor of most routine test environments may lead to invalidity at levels as low as 40 dB SPL.

In terms of clinical application, it is perhaps unfortunate that filter condition was not found to have a direct significant effect on CAEP detection, as this implies that the CAEP may not be sufficiently sensitive to very small changes in hearing aid characteristics to guide very fine-tuning, on the basis of current evidence. These small changes might even be more difficult to detect in infants because of their less reliable recordings and variable morphologies due to age and sound exposure (Sharma 2002). However, the stimulus sensation level (as determined behaviorally) was found to affect the CAEP, and more detailed investigation of responses at a wider range of consecutive sensation levels, may provide additional, clinically useful, information.

Results of this study showed that direct evaluation of audibility using CAEPs was slightly more precise than audibility prediction (based on threshold and hearing aid response measures), at least for adult listeners (refer to Figures 4 and 5). It is recognized that adults are more reliable subjects for CAEP assessment than infants, being less prone to inherent noise affecting the recording quality, among other factors. The estimation of hearing thresholds (as well as the recording of CAEPs) will be less precise for infants than in most adult cases. In addition, it is acknowledged that measurement of hearing aid gain will also be less accurate for hearing aids that incorporate more complex signal processing than was used in this experiment. Consequently, predicting audibility will be more difficult with hearing aids that incorporate more complex signal processing. However, it should also be considered that whatever the complexity of signal processing, hearing aids still function primarily to output speech sounds, and there is no reason to expect this to change markedly because of how a hearing aid algorithm determines the type of amplification provided – the important matter is the resulting behavioral sensation level of the speech sound received by the hearing aid wearer.

This experiment has clearly demonstrated that the presence of cortical responses is strongly related to the behavioral sensation level of speech sounds, which is encouraging in regards to the clinical application of CAEPs for hearing aid evaluation in the "difficult to test", including infants. However, it is important to keep in mind that at this stage the CAEP cannot add insight into the question of over-amplification, nor the verification of the maximum power output of the aid. In addition, there is still much to learn about prescribing optimal amplification characteristics for children in general. Whereas there is direct evidence to support the use of NAL-NL1 for adults, the evidence from children is less direct and more uncertain, and currently there is no evidence to support an intentional variation from the shape prescribed for an adult with the same hearing loss. Unfortunately, neither is there evidence that the same response should be prescribed (Dillon, 2012). Further, the uncertainties are greater for children with severe hearing losses than those with milder losses (Ching et al, 2002).

The concerns of previous authors regarding the current limits of our understanding of CAEPs (e.g., Billings et al, 2011) are acknowledged. However, it is important to consider that a lack of objective information regarding audibility of speech sounds and/or uncertainty about the appropriateness of an early fitting of an infant or "difficult to test" hearing aid candidate, can result in cautiousness on the part of the clinician responsible for prescribing hearing aids. This may result in the fitting being too conservative - that is the gain and/or maximum power output of the hearing aid being significantly under-prescribed, and/or delays in the decision to consider alternative strategies to amplification, such as augmented communication methods and/or cochlear implantation. This can deprive a child of language/auditory experience, particularly critical during the first year of life, when central pathways are being developed (Dillon et al, 2008). There may also be adverse effects arising from clinically conservative decision-making for an older person who is unable to communicate about his/her hearing aid fitting, and likewise needs optimal sensory input, in order to enhance quality-of-life. Conversely, when there is a lack of observed response to sound, the absence of objective information about audibility may also contribute to incorrect clinical interpretations in the opposite direction. In the event that an individual is actually able to hear sound, but is unable to show any behavioral response, an incorrect subjective judgement on the part of the clinician may result in over-prescription of gain and/or maximum power output, or even to hearing aids being fitted when they are not indicated.

The real possibility of non-detected corticals, however, must be kept in mind. In adults, some studies reported deviations as large as 30 dB between cortical and behavioral thresholds (Hoth, 1993) or up to 14.5% poorer CAEP thresholds than behavioral thresholds by 15 dB or more (Tsui, 2002) when presenting tone-bursts. In infants, this effect is even worse. Van Dun et al (2012) showed that for speech sounds presented in free field, only 75% of the CAEPs were detected for behavioral sensation levels above 10 dB SL. Such error may have serious and adverse ramifications in terms of causing discomfort to the individual (which they may not be able to communicate to the clinician or carer), and/or potential damage to residual hearing in the longer term. In addition, it is not currently possible to determine ‘adequate’ audibility for sufficient speech understanding on the basis of CAEPs, as these electrophysiological measures are mainly used to determine sound detection, as supported by this study.

CAEP testing for hearing aid evaluation can, however, be feasibly taken up in the clinical management of infants and difficult-to-test children. The authors note that currently (2012) in Australia almost 20 pediatric audiological centres use detection of CAEPs to speech sounds with different frequency emphases in free field to fine-tune hearing aid fittings, or to strengthen candidacy cases for cochlear implantation. According to a recent survey of clinicians conducted by the authors, the tested population is mainly composed of pediatric clients being fitted for the first time, those attending follow-up appointment for fine-tuning, cochlear implant candidates, and children with multiple disabilities. Audiologists provided a positive to very positive evaluation of this newly introduced technique, and indicated objective evaluation supported or supplemented their own findings. Appointments generally could be completed within 90 minutes, including hearing aid checks, setup, CAEP recording, and evaluation of the results.

CONCLUSIONS

Stapells (2002) commented that "the [cortical] response can provide functional measures of hearing aid benefit" and also notes that "the CAEP is the electrophysiological "measure of choice" when an estimate of hearing threshold is required for any patient who is likely to be passively (or actively) cooperative ("passively cooperative" includes watching TV, reading a book, even quiet play)". Although the focus of this study was the validity of the aided CAEP, the usefulness of unaided CAEPs should not be overlooked, particularly in the "difficult to test" population.

The current authors suggest that while there is still much to learn about the nature of the CAEP when hearing aids (and also cochlear implants) are worn, CAEP assessment of the young or "difficult to test" in the context of hearing aid evaluation, should not cause harm in itself and should not be avoided because of theoretical uncertainties. Of course, proper interpretation of results, and a clear understanding of the current knowledge of the characteristics of the CAEP, is the key to avoiding inappropriate clinical actions based on test results. For example, as emphasized previously, it is evident that a small proportion of children do not have observable CAEPs until the stimulus reaches a very high behavioral sensation level. The same seemingly applies to adults, and the reasons for this are not yet understood. It is therefore important to understand that the absence of a cortical response does not absolutely imply that the sound has not been detected by the hearing aid wearer (Dillon, 2012). Results must always be viewed in the context of the complete audiological assessment battery, including parent or habilitationalist reports of functional auditory behaviour.

When aided CAEPs are detected under good recording conditions this should, at least, provide reassurance that the stimuli of interest are audible to the hearing aid wearer (adult or infant). In this respect, CAEPs can already play a valuable role in counseling the parents of infants fitted with hearing aids, or carers of adults with developmental or other disabilities. The appearance of 'brain-wave' activity in response to conversational speech sounds, while wearing hearing aids, can reassure parents and/or professionals about the auditory capability of the individual, and reinforce to them the importance of the hearing aids being worn (Dillon, 2012). This demonstration is particularly salient when, in the same individual, CAEPs are not detected under the unaided condition. In the case of older "difficult to test" hearing aid candidates the information provided by CAEP assessment may be substantially more than could be gleaned with the traditional battery of hearing assessment and hearing aid evaluation methods, particularly for individuals that cannot undergo sedation to enable other forms of electrophysiological assessment. Even basic information regarding audibility of speech sounds has much to offer the audiologist in guiding fundamental clinical decisions such as whether or not amplification should be trialed, modified or withdrawn, and may potentially provide the opportunity for greatly improved outcomes for individuals unable to communicate directly about their auditory experience.

It is hoped that the results of this study will increase confidence in the use of CAEPs as a clinical tool, particularly for those clients for whom there is little reliable behavioral information regarding their hearing aid fitting outcomes. Ongoing research will hopefully provide further insights into the characteristics of the CAEP for children and adults wearing hearing aids under normal conditions of use (i.e., with advanced signal processing and features active).

REFERENCES

ABS (Australian Bureau of Statistics) (1993a) Summary of findings in, *Disability, ageing and carers, data reference package.* Catalogue number 4432.0 Canberra: AGPS.

ABS (Australian Bureau of Statistics) (1993b) *Disability, ageing and carers Australia: Brain injury and stroke.* Catalogue number 4437.0 Canberra: AGPS.

AIHW (Australian Institute of Health and Welfare) (2007) *Dementia in Australia: national data analysis and development*. Catalogue number AGE 53 Canberra: AIHW.

Agung, K, Purdy, SC, McMahon CM, Newall, P. (2006) The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *J Am Acad Audiol* 17(8): 559-572.

Allen, MC. (2008) Neurodevelopmental outcomes of preterm infants. *Curr Opin Neurol* 21: 123-128.

Alvarenga KF, Amorim RB, Agostinho-Pesse RS, Costa OA, Nascimento LT, Bevilacqua MC. (2012) Speech perception and cortical auditory evoked potentials in cochlear implant users with auditory neuropathy spectrum disorders. *Int J Pediatr Otorhinolaryngol* 76(9):1332-8.

Bentler, RA, Pavlovic, CV. (1989). Transfer functions and correction factors used in hearing aid evaluation and research. *Ear Hear* 10(1): 58-63.

Billings, CJ, Tremblay, KL, Miller, CW. (2011). Aided cortical auditory evoked potentials in response to changes in hearing aid gain. *Int J Audiol* 50(7): 459-467.

Billings C, Tremblay K, Stecker G, Tolin W. (2009) Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hear Res*.254(1-2):15-24

Billings, CJ, Tremblay, KL, Souza, PE, Binns, MA. (2007) Effects of hearing aid amplification and stimulus intensity on cortical auditory evoked potentials. *Audiol Neurootol* 12: 234-246.

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

Ching, TYC, Britton, L, Dillon, H, Agung, K. (2002) NAL-N1, RECD, & REAG: Accurate and practical methods for fitting non-linear hearing aids to infants and children. *Hear Rev* 9(8): 12-20, 52.

Ching, TYC, Dillon, H., Katsch, R, Byrne, D. (2001) Maximizing effective audibility in hearing aid fitting. *Ear Hear* 22(3): 212-224.

Cone-Wesson, B, Wunderlich, J. (2003) Auditory evoked potentials from the cortex: audiology applications. *Curr Opin Otolaryngol Head Neck Surg* 11(5): 372-377.

Dillon, H. (2001) *Hearing Aids*. Sydney: Boomerang Press.

Dillon, H. (2005). So, baby, how does it sound? Cortical assessment of infants with hearing aids. *Hear J* 58(1):10,12-17.

Dillon, H. (2012) *Hearing aids, 2nd edition*. Sydney: Boomerang Press

Dillon, H, Ching, T, Golding M. (2008) Hearing aids for infants and children. Chapter 18 in; Madell, JR, and Flexer, C. *Pediatric audiology, diagnosis, technology and management*. New York: Thieme.

Durrleman, S, Simon, R. (1989) Flexible regression models with cubic splines. *Stat Med* 8: 551-561.

Erber, NP, Witt, LH. (1977) Effects of stimulus intensity on speech perception by deaf children. *J Speech Hear Disord* 42 (2): 271-278.

Flury, B. and Riedwyl, H. (1988). *Multivariate Statistics: A practical approach*. London: Chapman and Hall.

Fortnum, H, Marshall, DH, Summerfield, AQ. (2002) Epidemiology of the UK population of hearing impaired children, including characteristics of those with and without cochlear implants – audiology, aetiology, co-morbidity and affluence. *Int J Audiol* 41(3): 170-179.

Frye, K, and Martin, R. (2008) *Audiology treatment, 2nd edition*. Valente, M, Hosford-Dunn, H, Roeser, RJ. Thieme, New York.

Garinis, AC, Cone-Wesson, BK, (2007) Effects of stimulus level on cortical auditory event-related potentials evoked by speech. *J Am Acad Audiol* 18: 107-116.

Golding M, Dillon H, Seymour J, Carter L. (2009) The detection of adult cortical auditory evoked potentials (CAEPs) using an automated statistic and visual detection. *Int J Audiol* 48(12): 833-42.

Golding, M, Pearce, W, Seymour, J, Cooper, A, Ching, T, & Dillon, H. (2007) The relationship between obligatory cortical auditory evoked potentials (CAEPs) and functional measures in young infants. *J Am Acad Audiol* 18: 117-125.

Golding, M, Purdy, SC, Sharma, M, Dillon, H. (2006) The effect of stimulus duration and inter-stimulus interval on cortical responses in infants. *Aust NZ J Audiol* 28 (2): 122-136.

Harris, RJ. (2001) *A primer of multivariate statistics*, 3rd edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Hoth, S. (1993). Computer-aided hearing threshold determination from cortical auditory evoked potentials. *Scand Audiol* 22, 165-177.

Hyde, M. (1997) The N1 response and its applications. *Audiol Neurootol* 2:281-307.

Jacobson G, Lombardi D, Gibbens N, Ahmad B, Newman C. (1992) The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component N1. *Ear Hear* 13(5):300-6.

Korczak, PA, Kurtzberg, D, Stapells, DR. (2005) Effects of sensorineural hearing loss and personal hearing aids on cortical event-related potential and behavioral measures of speech sound processing. *Ear Hear* 26 (2): 165-185.

Kurtzberg, D. Stapells, D, Wallace, IF. (1988) Event-related potential assessment of auditory system integrity: Implications for language development. Chapter 9 in; Vietze, PM. Vaughn, HR. *Early identification of infants with developmental disabilities*. Philadelphia: Grune and Stratton.

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

Mace, AL, Wallace, KL, Whan, MQ, Stelmachowicz, PG. (1991) Relevant factors in the identification of hearing loss. *Ear Hear* 12(4): 287-293.

Martin, B., Tremblay, KL., Stapells, DR. (2007). Principles and applications of cortical auditory evoked potentials. Chapter 23 in; *Auditory evoked potentials basic principles and applications*. R. Burkard, Don. M., Eggermont, JJ., Lippicott, Williams & Wilkins: Pages 497-507.

Mendel M, Hosick E, Windman T, Davis H, Hirsh S, Dinges D. (1975) Audiometric comparison of the middle and late components of the adult auditory evoked potentials awake and asleep. *Electroencephalogr Clin Neurophysiol*. 38(1):27-33

Moore, BCJ, Glasberg, BR, (2004). A revised model of loudness perception applied to cochlear hearing loss. *Hear Res* 188: 70-88.

Palmer, C. (2005) In fitting kids with hearing aids, ensuring safety and audibility is a good way to start. *Hear J* 58(2): 10-16.

Picton TW. (2011) *Human auditory evoked potentials* San Diego, CA: Plural Publishing Inc.

Picton, T., Hillyard, S. (1974). Human auditory evoked potentials. II: Effects of attention. *Electroencephalogr Clin Neurophysiol* 36: 191-200.

Picton T, Woods D, Baribeau-Braun J, Healey T. (1977) Evoked potential audiometry. *J Otolaryngol.* 6:90-119.

Purdy, SC, Katsch, R, Dillon, H, Storey, L, Sharma, M, Agung, K. (2005). Aided cortical auditory evoked potentials for hearing instrument evaluation in infants. Chapter 8 in; Seewald, R. (Ed) *A sound foundation through early amplification. Proceedings of the third international conference.* Warrenville, IL: Phonak.

Purdy, SC, Kelly, AS. (2001). Cortical auditory evoked potential testing in infants and young children. *NZ Audiol Soc Bull* 11(3):16-24.

Ray, C. (2002) Mental retardation and/or developmental disabilities. Chapter 30 in; Katz, J. (Ed) *Handbook of clinical audiology, 5th edition.* Baltimore, Philadelphia: Lippincott, Williams & Wilkins.

Rickards, F., De Visi, S, McMahon, DS. (1996). Cortical evoked response audiometry in noise-induced hearing loss claims. *Aust Otolaryngol* 2: 237-241.

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

Runge C, Gilham J, Peut, A. (2009) *Transitions in care of people with dementia*. Catalogue number AIHW 11074. Canberra: AIHW.

Scollie, SD, Seewald, RC. (2002) Electroacoustic Verification measures with modern hearing instrument technology. In: Seewald, RC, Gravel, JS (Eds) *A sound foundation through early amplification.* 2001 Phonak AG.

Seewald, RC, Moodie, S, Scollie, S, Bagatto, M. (2005) The DSL method for pediatric hearing instrument fitting: Historical perspective and current issues. *Trends Amplif* 9(4): 145-157.

Sharma A, Dorman M, Spahr A. (2002) A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear* 23(6):532-9.

Sharma A, Martin K, Roland P, Bauer P, Sweeney MH, Gilley P, et al. (2005) P1 latency as a biomarker for central auditory development in children with hearing impairment. *J Am Acad Audiol*. 16(8):564-73.

Souza, PE, Tremblay, KL. (2006) New perspectives on assessing amplification effects. *Trends Amplif* 10(3): 119-143.

Stapells, DR Cortical event-related potentials to auditory stimuli. Chapter 20 in; Katz, J, (2002) *Handbook of clinical audiolology, 5th edition*. Philadelphia, PA: Lippincott, Williams & Wilkins.

Stelmachowicz, P, Hoover, B. (2009) In; Katz, J, Medwetsky, L, Burkard, R, Hood, L. *Handbook of clinical audiology, 6th edition*. Baltimore, Maryland; Philadelphia, PA: Lippincott, Williams & Wilkins.

Stelmachowicz, PG, Hoover, BM, Lewis, DE, Kortekaas, RWL, Pittman, AL. (2000) The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *J Speech, Lang and Hear Res* 43 (4): 902-914.

Thabet MT, Said NM. (2012) Cortical auditory evoked potential (P1): A potential objective indicator for auditory rehabilitation outcome. *Int J Pediatr Otorhinolaryngol* 2012 Aug 31. [Epub ahead of print]

Thompson, G, and Weber, BA. (1974) Responses of infants and young children to behavior observation audiometry (BOA). *J Speech Hear Disord*. 1974 39(2): 140-7.

Tremblay, KL, Billings, CJ, Friesen, LM, Souza, P. (2006a) Neural representation of amplified speech. *Ear Hear* 27(2): 93-103.

Tremblay, KL, Kalstein, L, Billings, CJ, Souza, PE. (2006b) The neural representation of consonant-vowel transitions in adults who wear hearing aids. *Trends Amplif* 10(3): 155-162.

Tsui, B., Wong, L., Wong, E. (2002). Accuracy of cortical evoked response audiometry in the identification of non-organic hearing loss. *Int J Audiol* 41(6): 330-333.

Van Dun B, Carter L, Dillon H. (2012) Sensitivity of cortical auditory evoked potential detection for hearing-impaired infants in response to short speech sounds. *Audiol Res* 2e13: 65-76.

Wen, X. (2007) *The definition and prevalence of intellectual disability in Australia*. Australian Institute of Health and Welfare, Canberra. AIHW catalogue number DIS 2.

Wunderlich J, Cone-Wesson B, Shepherd R. (2006) Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res* 212(1-2):185-202.

Title: Cortical Auditory Evoked Potentials (CAEPs) in adults in response to filtered speech stimuli.

Authors: Lyndal Carter, Harvey Dillon, John Seymour, Mark Seeto, Bram Van Dun

FIGURE LEGEND

TABLE 1= Nominal test runs at each calculated sensation level range by stimulus.

TABLE 2= Percentage of CAEPs with detection *p*-value < 0.05 for all speech stimuli by behavioral sensation level range.

TABLE 3= Latency difference estimates (in ms) and their 95% confidence intervals for stimulus and filter (stimulus reference = /m/, filter reference= low-boost). Negative values indicate that variable latencies are shorter than their reference.

FIGURE 1= Pure tone audiometric thresholds for test ear of each participant.

FIGURE 2= Test device fit to NAL-RP target (REIG – NAL-RP).

FIGURE 3= Calculated SL versus behaviorally measured SL (all stimuli and conditions).

FIGURE 4= *p*-value (Hotelling's T^2) as a function of behavioral SL.

FIGURE 5= *p*-value (Hotelling's T^2) as a function of calculated SL.

FIGURE 6= *p*-value (Hotelling's T^2) as a function of 4FAHL.

FIGURE 7= Grand average waveform for SL > 10 dB: all conditions, all participants.

FIGURE 8= $P_2 – N_1$ amplitude versus 4FAHL (test ear).

FIGURE 9= $P_2 – N_1$ amplitude versus behavioral SL on CAEP with fitted curves for each stimulus.

FIGURE 10= Histogram of residual noise (adjusted for number of epochs).

FIGURE 11= Residual noise (adjusted for number of epochs), by 4FAHL.