



National
Acoustic
Laboratories

Report: Clinical Evidence for HEARLab
(Cortical auditory evoked response testing modules)

Authors: Maryanne Golding, Ph.D
Lyndal Carter, M.Aud

Date: 2nd October, 2007

Summary	3
Objectives	4
Part A: Literature review	5
Part B: Clinical verification	18
1. Method	19
1.1 Participants.....	19
1.2 Stimuli.....	20
1.3 Procedure	20
1.4 Response detection and differentiation.....	23
2. Results.....	24
2.1 Latency.....	24
2.2 Amplitude	26
2.3 Detection P values.....	29
2.4 Difference P values	34
2.5 Reject rate	36
2.6 Impedance	37
2.7 Residual noise (“RN”)	39
3. Discussion.....	41
3.1 Latency.....	41
3.2 Amplitude	42
3.3 Detection P values.....	43
3.4 Difference P values	46
3.5 Reject rates and residual noise.....	46
4. Conclusion and recommendations	47
Reference List	48
Appendix 1: Record CAEPs in infants with success	55
Appendix 2: Cochlear implants and HEARlab	63

Summary

The overall aim of this study was to a) examine the suitability of HEARLab's selected test parameters and user-interface for CAEP testing, and b) perform a clinical study in which the accuracy of HEARLab's two modules in recording CAEPs was compared with responses recorded simultaneously from a second device (i.e., Neuroscan TM) that is known to produce true electrophysiological outcomes.

The majority of components within HEARLab's device requirements specifications were justifiable and require no further action. Some minor exceptions were found and reported. The clinical study showed few system differences of note. One exception was a significant difference in the statistical detection P value measure between systems where HEARLab was found to be more conservative than Neuroscan TM in detecting CAEPs. One manifestation of this was reported as clinically insignificant but the second was related to the rate of absent cortical responses and as such needs to be addressed. Recommendations are made on how the source of this finding might be investigated and identified.

Objectives

The first module of the HEARLab instrument is designed to record obligatory cortical auditory evoked potentials (CAEPs) (i.e., cortically-generated responses to any audible stimulus of an appropriate length) in response to speech and tonal stimuli. A clinical verification of the device needs to be completed as a requirement for medical device regulatory authorities before commercial release.

An extensive literature review will be performed that will examine the suitability of HEARLab's selected test parameters and user – interface for CAEP testing. Evidence for the accuracy of the modules (i.e., Aided Cortical Assessment (ACA) and Cortical Threshold Evaluation (CTE) modules) in recording CAEPs will be gathered from a clinical study in which these responses will be generated from adult and infant participants. These responses will be concurrently recorded from HEARLab and a second device known as Neuroscan™ that is known to produce true electrophysiological outcomes. The study will compare:

- a) electrode impedance,
- b) rejection rate of unacceptable epochs,
- c) accuracy of response amplitude and latency measures,
- d) statistical outcomes (i.e., detection p values and difference p values),
- e) test “power” measures,

across the two devices.

As a secondary objective, HEARLab's reliability will also be examined using repeat assessments in adult participants only, as it is unlikely that infants can be recruited to attend on two occasions. The reliability of any device, however, depends on the stability of the attribute under measure across time. It is expected that the presence/absence of the cortical response, should be stable across visits provided that a similar signal-to-noise ratio (SNR) can be achieved, the participant is equally alert on both occasions and has not suffered any trauma to their hearing between visits.

Part A: Literature review

The literature review has been prepared in conjunction with the “Software Requirement Specification Guide: device requirements specification”. Each component of these specifications has been examined and reported against in terms of “action” and “comments”.

SRS Clause	Action	Comments
3.1 Device primary functional requirements		
3.1.1 Presents acoustic stimulus		
3.1.1.1 Types of auditory signals	Closed	Appropriate
3.1.1.1.1 Speech sounds	Use modified stimuli	Three speech stimuli (/m/, /t/, /g/) are available through the ACA unit. These stimuli were originally generated from natural speech tokens consisting of an initial consonant followed by the vowel /ae/ which was extracted from a recording of running speech that was spoken by a male with an average Australian accent. The final test stimuli included very little of the vowel transition and were recorded with digitization rates of 40 kHz. They were gated off near/at a zero crossing to avoid audible clicks. These essentially vowel-free stimuli were chosen because they had a spectral emphasis in the low-, mid- and high-frequency regions respectively and thus had the potential to give diagnostic information about the perception of speech sounds in different frequency regions. They have been extensively used in cortical response projects at NAL (Dillon, Golding, Purdy, & Katsch, 2006; Golding, Dillon, Seymour, Purdy, & Katsch, 2006; Golding, Pearce, Seymour, Cooper & Ching, 2007) for the assessment of infants with normal hearing and those fitted with hearing aids. The current modified speech tokens, /m/, /g/ and /t/ were extracted from ILTASS-filtered continuous discourse spoken by a female with an average Australian accent and /t/and /g/ were high passed filtered at 250 Hz. These modified stimuli show more distinctive peak energy within the low-, mid- and high- frequency bands than the original stimuli. Adult-generated cortical responses recorded to the original and modified stimuli were compared and no detrimental differences were found when the modified stimuli were used (Golding, 2006). HEARLab ACA is an implementation of NAL research findings for a clinical setting.

3.1.1.1.2 Duration of speech sounds	Use modified stimuli	In a NAL study using the original /m/ and /t/ stimuli, the effect of differing stimulus durations over a limited range (i.e., /m/ at 31 ms, 79 ms and 141 ms; and /t/ at 32 ms and 79 ms) were examined. Golding et al. (2006) reported that in infants, there was no significant simple effect of duration on either the amplitude or latency of the cortical responses for /m/ or /t/. When data from /m/ and /t/ stimulation were combined and responses to short (31-32 ms) versus medium (79 ms) duration stimuli were compared, there was some amplitude increase from short to medium durations. The current modified speech tokens are shorter than the original (/m/: original =79 ms, modified = 30 ms; /g/: original = 31 ms, modified = 20 ms; /t/: original =79 ms, modified = 30 ms) but as reported above and under 3.1.1.1.1, no simple effects of duration (within the range examined) on amplitude or latency have been found. The duration of the modified stimuli is therefore acceptable.
3.1.1.1.3 Types of transducers	Closed	HEARLab presents speech sound test signals in the free-field only. This is appropriate for the evaluation of hearing pre- and post fitting of hearing aids. Distances between the participant's head and the loudspeaker in the free-field have varied across studies from 7.6 cm (Brown, Klein, & Snyder, 1999), 20 cm (Kushnerenko et al., 2002) to 1 meter (Simos & Molfese, 1997). The angle between the head and the speaker is not always reported in studies but when it is noted, it varies significantly (Kurtzberg, 1989; Kushnerenko et al., 2002). The use of a loudspeaker positioned at 90° azimuth relative to the test ear (TE) has been reported (Brown et al., 1999) as has the use of two loudspeakers positioned at 45° facing the right and left ears (Purdy et al., 2005). One study has reported the use of a loudspeaker positioned directly above the participant and equidistant to both ears (Simos et al., 1997). It can be concluded therefore that while it is important to measure the stimulus intensity at the TE, the issues of angle to the speaker and distance from the participant are not critical to the recording of the cortical responses. It is likely that practical considerations associated with the size of the test room will be the likely determinant of the position of the speaker(s) relative to the participant. It is important to note however that the acoustic calibration of the system should be performed using the same spatial arrangements between the participant's head and loudspeaker as that intended for testing.

3.1.1.1.4 Stimuli presentation level	Closed	The ACA module of HEARlab offers a limited stimulus presentation level range of 55 – 75 dBSPLi. This is appropriate as the intended clinical application for this unit is to assess an infant’s ability to detect speech stimuli presented at conversational levels. CAEPs can be observed reliably to speech stimuli presented at these levels in normal hearing infants (Cone-Wesson & Wunderlich, 2003; Kurtzberg, 1989)) and a relationship between the presence/absence of aided cortical responses to functional performance has also been demonstrated in infants (Golding et al., 2007). Using this technique to estimate auditory threshold in infants has however not been established reliably.
3.1.1.1.5 Polarity	Closed	Speech stimuli are presented with alternating polarity only in the ACA unit. The selection of signal polarity is not critical to recording CAEPs (Hall, 2007; Hyde, 1997) but as the use of stimuli that alternate in polarity is known to be effective in reducing the contaminating effects of stimulus artifact upon the waveform (Hall, 2007; Hyde, 1997; Goldstein & Aldrich, 1999), it has been incorporated in this module.
3.1.1.1.6 Repetition rates	Closed	Speech sound test signals are presented in the ACA unit with an inter stimulus interval (ISI) (i.e., the period between stimulus offset and the following stimulus onset) of 1125ms. In studies where adult-generated cortical responses are recorded, an ISI of 1 or 2 seconds has been reported as clinically satisfactory (Hyde, 1997; Stapells, 2002). Variations in the ISI with the speech stimuli /uh/, have been recently shown to impact on the components of the cortical response in children 3-12 years in a highly complex manner (Gilley, Sharma, Dorman, & Martin, 2005) but the impact of ISI change in infants has not been widely reported. In a NAL study using the original /m/ and /t/ stimuli, the stimulus duration was fixed at 79 ms but the ISI was varied from 750 ms to 1125 ms and 1500 ms. In infants, the amplitude of the cortical response increased as the ISI increased from 750 to 1500 ms for stimulus /t/ only but there was no difference between an ISI of 1125 ms and 1500 ms (Golding et al., 2006). As there was no clear advantage in using the 1500 ms rate and test time would be increased, the shorter 1125 ms ISI appears to be a satisfactory and appropriate interval for all stimuli.

3.1.1.1.7 Masking	Closed	Masking is not provided in the ACA module. If the clinician wishes to evaluate each aided ear individually, testing can be performed with the non-test ear (NTE) occluded by the child's own earmould and hearing aid in a switched-off position (Golding et al., 2007). For unaided testing, individual ear responses are unlikely to be necessary but ear plugs could be inserted if necessary.
3.1.1.2 Tone-bursts	Closed	An appropriate range of tone-burst frequencies are available for threshold estimation through the CTE module. This module is likely to be used for auditory threshold estimation in adults and older children when standard behavioural techniques are ineffective (e.g., noise-induced hearing loss compensation assessments). The presentation frequencies offered through CTE are 500, 1000, 1500, 2000, 3000 and 4000 Hz. This range is consistent with conventional audiometry although 250 and 8000 Hz are not offered. While it could be argued that these extreme audiometric frequencies should be available through the CTE module for diagnostic purposes (e.g., Meniere's disease, ototoxicity), responses to 250 and 8000 Hz are least likely to influence rehabilitation outcomes, and they are not required for the calculation of percentage hearing loss (NAL report, 1982).
3.1.1.2.1 Tone-burst characteristics	Closed	The tone-burst envelope comprises a rise and fall time of 10ms each and a plateau of 30ms. The rise and fall shape follows the cosine curve over 90° with an equivalent overall period of 40ms. Variations to the rise/fall and the duration of tonal stimuli have complex implications for the amplitude and latency of adult cortical responses. When the rise/fall is brief (e.g., 3 ms), an increase in response amplitude and a decrease in response latency might be expected as the duration of the stimulus plateau varies from 0 – 30 ms and up to 70 ms (Alain, Woods, & Covarrubias, 1997; Onishi & Davis, 1968). When the rise/fall is 5 ms, the amplitude of the cortical response is nearly constant as the duration is varied from 2 – 320 ms (Davis & Zerlin, 1965)). Optimal responses are however recorded with rise/fall times and plateau times of greater than 10 ms (Onishi et al., 1968; Rothman, Davis, & Hay, 1970). The stimulus envelope offered in the CTE module is therefore selected to meet this requirement but it is not long enough to overlap with the response. As a result, the stimulus is fully effective in influencing the response outcome.

3.1.1.2.2 Centre-frequency of tone-bursts in Hertz	Closed	CTE module offers tone-bursts centered at the frequencies of 500, 1000, 1500, 2000, 3000 and 4000 Hz which is appropriate for estimation of auditory threshold across the standard audiometric range.
3.1.1.2.3 Types of transducers	Closed	<p>The CTE module delivers stimuli by insert earphones and bone conduction (BC) only.</p> <p>Insert earphones offer a number of advantages over headphones that apply equally to evoked potential testing and standard audiometry. Firstly, they offer increased accuracy in the estimation of threshold in the low frequencies because of reduced variability arising from air leaks around standard headphones (Gordon, Phillips, Helt, Konrad-Martin, & Fausti, 2005; Zwislocki et al., 1988). Secondly, the insertion depth that is required for positioning the device attenuates physiological noise in the ear and external ambient noise which makes the use of sound-treated test facilities less critical (Clemis, Ballard, & Killion, 1986; Gordon et al., 2005; Zwislocki et al., 1988). Thirdly, the potential for ear canal collapse in older adults, and hence false air conduction (AC) thresholds, is eliminated (Clemis et al., 1986; Gordon et al., 2005). Fourthly, insert earphones maximize inter-aural attenuation and reduce the need for masking in the NTE which is particularly useful when significant asymmetric or conductive hearing loss exists (Clemis et al., 1986; Gordon et al., 2005; Killion & Villchur, 1989; Zwislocki et al., 1988). Finally, they provide increased comfort when worn for extended periods (Clemis et al., 1986; Gordon et al., 2005).</p> <p>The recording of cortical responses to bone conducted tone-burst stimuli is not performed routinely but its application should not be considered any more problematic than in behavioural testing (Durrant & Hyre, 1993). For unmasked bone conducted responses, placement of the device on the electrode-free mastoid (see 3.1.2.2) should be satisfactory given that the head offers very little attenuation to BC stimulation in adult subjects (Hall, 1992). If time permitted, the examiner could find estimates of bone conducted thresholds monaurally using the appropriate application of masking to the NTE (see 3.1.1.2.7 for comments on masking). The application of the “reference” electrode to the NTE would be appropriate in these circumstances (see 3.1.2.2 for comments on electrode placement)</p>

3.1.1.2.4 Stimuli presentation level	Closed	The CTE module offers a stimulus presentation range from 0 – 110 dB HL via insert earphones and 0 – 70 dB HL via BC. The examiner must choose the stimulus presentation level before selecting the test frequency which is counterintuitive but satisfactory. The step-size for both AC and BC presentations is 5 dB which is consistent with standard audiometry. The relationship between electrophysiological and behavioural threshold can therefore be readily explored.
3.1.1.2.5 Polarity	Closed	The tone-burst stimuli within the CTE module are presented with an alternating polarity which is appropriate (see comments 3.1.1.1.5)
3.1.1.2.6 Repetition rates	Closed	The ISI for tone-burst stimuli within the CTE module is 1125 ms. This is well within the recommended range of one to two per second (Abramovic, 1990) that optimizes the amplitude of the cortical response (Goldstein et al., 1999). It is consistent with the rate offered for the presentation of speech stimuli within the ACA module.
3.1.1.2.7 Masking	Modify masking level range	<p>Narrow-band noise is offered as the masking signal in the CTE module.</p> <p>Narrow bands of noise centred at pure-tone test frequencies are routinely and effectively applied to the NTE in audiometry (Hall, 2007; Hyde, 1997) and are appropriate for cortical testing when tone burst stimuli are delivered via AC or BC (Goldstein et al., 1999). The level of masking applied to the NTE, may follow normal audiometric rules (Reid & Thornton, 1983). This would involve unmasked threshold estimation for all frequencies of interest in both ears before re-estimating thresholds (with masking applied to the NTE) for those frequencies where a cross-heard stimulus is possible. Given the increased inter-aural attenuation offered by insert earphones, masking to the NTE may not be needed for stimulus levels in the TE \leq 70 dB HL as inter-aural attenuation for insert earphones may be as much as 60 dB (Hall, 2007). For stimulus presentation levels above this (or when stimulation is by BC), the CTE module offers a masking level range of +10 to -30 dB (relative to the stimulus presentation level). With an extreme stimulus presentation in the TE of 110 dB, the contralateral masking will be 80 dB or more which may</p>

		lead to over-masking in some cases. An increase in masking level range is recommended (i.e., recommend +10 to -60 dB relative to the stimulus presentation level).
3.1.2 Acquire acoustically evoked responses		
3.1.2.1 No. of signal channels and electrodes	Closed	<p>The ACA and CTE modules offer a single channel for signal detection.</p> <p>For cortical response detection a single channel is quite satisfactory (Roger & Thornton, 2007) and given the practical challenges of attaching electrodes and maintaining placement during testing, HEARLab's single channel recording system is appropriate. In a recent NAL study where the validity of a statistical technique in the detection of cortical responses was evaluated (see 3.1.4.1), cortical responses from ten adult subjects were recorded to two speech stimuli presented at five sensation levels (relative to their auditory threshold). While recordings were made at three electrode sites; C3, Cz, C4, no discernable difference in response morphology could be observed across the three sites (Golding, Dillon, Seymour & Carter, 2007). As a result reporting was restricted to responses recorded using the single channel of Cz.</p>
3.1.2.2 Electrode placement	Closed	<p>For ACA and CTE modules, the examiner is advised to position the active electrode on the vertex (Cz*), the reference electrode on the earlobe (A1 or A2*) or on the mastoid (M1 or M2*), and the ground electrode on the forehead (Fz*).</p> <p>For cortical response detection, placement of the active electrode at Cz or within two to three cm of the site ensures optimal amplitude of the cortical response in most cases (Hall, 2007). In recent NAL studies, where the effect of electrode montage on infant and adult cortical responses was specifically examined, the dominant adult N1 response was largest at Cz (compared with C3 and C4) (Purdy et al., 2006) while the dominant infant P1 was largest at C3 by 1 to 2 μV (Golding et al., 2006). A forehead position is a feasible alternative to placement at the vertex but an undesirable increase in eye movement artifact will be evident in the recording (Hyde, 1994).</p> <p>The use of mastoid or earlobe placement for the reference electrode is common (Goldstein et al., 1999;</p>

		<p>McPherson, 1996)but given the small surface area of an infant's earlobe, mastoid placement may be more practical in these cases. The detection of the cortical response is independent of the ear under stimulation and so it is not necessary to move the reference electrode from one mastoid to the other during testing (Stapells, 2002).</p> <p>Although many positions for the ground electrode are plausible, a forehead position is one of the easiest to achieve (Goldstein et al., 1999).</p> <p>** using the10-20 International Electrode (placement) system (Jasper, 1958).</p>
3.1.3 Display responses		
3.1.3.1 Display acquired signal	Show the X-axis parameter	<p>The CTE and ACA module show a continuous streaming of the electroencephalographic (eeg) signal. The amplitude of the in-coming signal is shown on the Y-axis while the X-axis is divided into 10 x 1 sec units although this parameter is not shown. A heavier vertical marker is also evident at regular intervals throughout the recording and is consistent with the stimulus trigger onset. Clinicians can monitor this acquired signal for signs of interference and other anomalies that are likely to result in a poor signal to noise ratio. The clinician may choose to cease testing until the source of these abnormalities can be identified and rectified.</p>
3.1.3.2 Display the response of the latest epoch	Closed	<p>The latest recorded epoch is displayed and can be viewed in conjunction with the display of the acquired signal as a means of monitoring the quality of the signal.</p> <p>Time (ms) is displayed on the X-axis as a function of amplitude (μV) on the Y-axis. These parameters are appropriate (see 3.1.3.3).</p>
3.1.3.3 Calculate and display average accumulated epochs	<p>Justify the absence of waveform from 550 – 600 ms.</p> <p>Increase the pre-stimulus region to -200ms.</p>	<p>The cumulative averaged cortical response is displayed in both the ACA and CTE modules. A shaded age-appropriate latency band for the first positive peak (mean P1 latency\pm 2 SD) has been added to the graph.</p> <p>The cumulative average response is shown in the time domain with latency (measured in ms) displayed on the X-axis as a function of amplitude (measured in μV) on the Y-axis. The latency scale is fixed at - 100 to + 600 ms (although waveforms cease at 550 ms) while the amplitude scale defaults to -15 to +15 μV but it can be modified by the examiner (unlimited scale).</p>

		<p>As cortical responses are particularly susceptible to changes in client state (Goldstein et al., 1999; Hall, 1992; Hyde, 1994; Stapells, 2002), the addition of a pre-stimulus baseline (-100 to 0 ms) is advantageous and its duration usually approximates 10% of the total analysis period (Hall, 1992). Baseline correction, which estimates the DC offset and then subtracts the average value from each sampling point within the epoch, allows the examiner to judge the activity level of the client simultaneously with the detection of the evoked response. This provides confirmation that the post-stimulus response is indeed stimulus-evoked and not simply a continuation of the pre-stimulus eeg. Given the low frequency band-width of the cortical response, a pre-stimulus region of -100 to 0 ms may not however be an adequate time period on which to base this estimate. An increase in the duration of the pre-stimulus region to -200 ms should be considered.</p> <p>The morphology of the major components of the cortical response change substantially with respect to the shape and latency over the first 14-16 years of life (NAL report 2005). The newborn infant cortical response is dominated by a series of positive peaks with a prominent peak at 200 to 300 ms when recorded at the midline (Kurtzberg, 1989; Sharma, Dorman, & Spahr, 2002; Stapells & Kurtzberg, 1991). By adult years (i.e., over 20 years), the dominant component is a negativity (80 – 120 ms) that is preceded and followed by positive components (i.e., P1 at 50 to 70 ms, and P2 at 150 – 200 ms) (Davis, 1965). The duration of the X-axis is therefore sufficient to accommodate these age-related changes to latency.</p> <p>The age-appropriate latency band was derived from a NAL study (Golding et al., 2006). The latency results for P1 in 54 infants, aged 0.2 to 0.75 years were plotted together with adult latency values and data published by Sharma and colleagues in which the P1 latency for adult normal hearers and those with cochlear implants is reported (Sharma et al., 2002). The shaded band that is superimposed on the averaged display is a true representation of these findings.</p>
--	--	--

3.1.4 Perform statistical analysis on acquired cortical responses		
3.1.4 .1 Detect acoustically evoked cortical response	Closed	<p>ACA and CTE modules apply Hotelling's T^2 (Flury & Riedwyl, 1988; Harris, 2001) for detection of cortical responses.</p> <p>To achieve this, accepted epochs to each polarity of each stimulus are paired and averaged. The sampling points for this averaged response are reduced to nine variables to form a "response" condition for further analysis. As the latency of key response components varies with age (see 3.1.3.3), the point of commencement and duration of the analysis period also varies. For infants 0 – 2 years, the analysis period commences at 101 ms (post-stimulus onset) with an end point at 550 ms; for 2 – 10 year olds, analysis commences at 76 ms (end point at 450 ms); for ages > 10 years, analysis commences at 51 ms (end point at 350 ms). With ten averaged responses per stimulus, Hotelling's T^2 is applied to calculate the probability that the mean value of any linear combination of the nine variables is significantly different from zero. The resulting P value is presented numerically and also graphically using a colour-code that is consistent with the cumulative average display. This value is continually refreshed with every increment of ten in averaged responses, until the stop-criteria for the test is reached.</p> <p>Automated and machine scoring methods for detection of evoked responses is not new but it has 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 infant-generated cortical responses (Golding et al., 2007). Its accuracy in detecting the presence of a cortical response when a stimulus was present and its accuracy in reporting the absence of a response when a non-stimulus trial was presented, has been compared with that of human examiners using infant- and adult-generated cortical responses. Results suggested that Hotelling's T^2 was at least equal to, if not better at detecting cortical responses than the average human expert (Golding et al., 2007).</p>

<p>3.1.4.2 Differentiate evoked cortical response to one stimulus to another stimulus</p>	<p>Closed</p>	<p>In the ACA module a Multivariate Analysis of Variance (MANOVA) is automatically applied to each stimulus pair (/m/ vs /t/, /t/ vs /g/, /m/ vs /g/) to establish whether a significant difference between response pairs can be detected. When ten averaged responses (see 3.1.4.1) for each stimulus within the pair are detected, two “response” conditions are formed (i.e., one in response to each stimulus under comparison) and these two conditions, nine variables (see 3.1.4.1) and ten epochs are used as input to the MANOVA from which a P value is reported numerically and graphically (see 3.1.4.1). This value is refreshed with every increment of ten averaged responses for each stimulus within the pair until the stop-criteria is reached.</p> <p>A recent study at NAL showed that 19 out of 20 infants with normal hearing had statistically different responses for /g/ versus /m/ and for /t/ versus /m/. Around 50% also showed a significantly different response for /t/ versus /g/ (Dillon, 2005; Purdy et al., 2005) . The relationship between differentiation of cortical responses and speech discrimination is, however, yet to be determined.</p>
<p>3.1.4.3 Perform a standard error (SE) measure and display</p>	<p>Closed</p>	<p>The ACA and CTE modules display an indicator of residual noise which is based on the inter-epoch variability and the number of accepted epochs in each averaged stimulus-generated response. This indicator provides the examiner with desirable information on the likelihood of detecting a response given the number of accepted epochs and electrical activity within the accumulated epochs.</p> <p>The RMS value of the residual noise that is superimposed on the averaged accepted epochs is calculated concurrent with the detection p value (see 3.1.4.1) and results are reported graphically. The colour “red” is used to denote a SE for the RMS of accepted epochs $> 4 \mu V$; “yellow” denotes $3 \mu V < SE < 4 \mu V$; “green” $< 3 \mu V$.</p> <p>The RMS and SE for accepted epochs was recently calculated for 183 cortical response eeg files from infants aged 6 months to 3.5 years (Golding, Dillon, & Seymour, 2006). These files all contained at least 100 accepted epochs. The SE was plotted as a function of the cortical grade (0 to 3) which was determined previously for these responses (Golding et al., 2007). There was consistency in the SE rate across cortical grades with the majority of data points contained within a band of 3 to $4 \mu V$. The cut-offs for the three categories were based on these findings.</p>

--	--	--

3.2 Other device functional requirements

3.2.1 Assessment recording keeping	Closed	It is necessary to store results electronically for retrieval at a later date. The ACA/CTE modules fulfil this requirement satisfactorily. Access to the store information is via the “client search function” where direct entry of the client’s unique ID or a “list all clients” option can be utilized. Each visit is stored by date and time and results can be easily retrieved and viewed. A number of “runs” may have been saved for each visit and/or assessment and so the clinician can choose which “run” to view.
3.2.2 Report	Closed	The provision of a printed form of the assessment allows the clinician to add a hard-copy of results to other client file notes if required. Reports can be printed in black and white or colour as required. This can be done at the time of testing or reports can be retrieved and printed at a later stage. The report screen contains the selected averaged response “run(s)”, a table of key wave component latency values as marked by the examiner, and detection p values. The ACA module report shows the detection p values and differentiation p values in separate tables. The client ID, date of test and session details are also provided. This is an appropriate summary of the test outcomes.

In summary, the majority of components within the device requirements specifications are justifiable and no further action is required. Exceptions are:

- a) 3.1.1.1.1 speech sounds,
- b) 3.1.1.1.2 duration of speech sounds,
- c) 3.1.1.2.7 masking,
- d) 3.1.3.1 display acquired signal,
- e) 3.1.3.3 calculate and display average accumulated epochs.

Part B: Clinical verification

Clinical evidence for the accuracy of the cortical response testing modules of HEARLab was gathered using concurrent recordings from the HEARLab and Neuroscan TM systems.

1. Method

1.1 Participants

Adult participants (N =12)

There were four males and eight females, with mean age of 44 years (SD 15), who completed this study. Of these 12 adults, six had pure tone thresholds in both ears \leq 20 dB HL for 500, 1000, 2000 and 4000 Hz. These six participants reported the absence of otological symptoms (i.e., tinnitus, otalgia and vertigo), and they had normal middle ear pressure and tympanic membrane compliance. The remaining six adults had a sensorineural impairment with a mean four frequency average (4FA) (500, 1000, 2000, 4000 Hz) in the TE of 37 dB HL (SD 17). Of these six participants, four regularly wore hearing aids. All six participants had normal middle ear pressure and tympanic membrane compliance.

A further two adults with cochlear implants volunteered to assist but were excluded as unaided threshold assessment would deliver limited data. They were however assessed in the free-field using the ACA module and these results are presented in Appendix 2.

Infant participants (N= 6)

Although 13 infants were recruited for this study, data was not obtained or excluded from analysis for seven infants. For three of these participants, experimental set-up problems occurred which could not be overcome within the time that the infant remained in a suitable state for testing. Another infant arrived in an inappropriate state for testing as they were unwell and became restless and overheated quickly. A further participant became distressed by the placement of the electrodes to the extent that testing could not be attempted. In the case of the final two participants, testing was abandoned before sufficient useful data could be obtained as the number of rejected epochs was very high which prolonged the length of the test session unacceptably.

Complete datasets were obtained on six infants (two males and four females) who ranged in age from 3 to 25 months. All were born at 37 – 40 weeks gestational age. While one of these six infants had a permanent hearing loss and had been fitted with binaural behind-the-ear hearing aids prior to this experiment, the remainder had “passed” neonatal automated auditory brainstem response testing and were not considered “at risk” for hearing loss. The parents of these five infants had no concerns about their child’s hearing ability in everyday situations.

1.2 Stimuli

Tone burst stimuli (see Part A 3.1.1.2 for complete description) with centre frequencies of 500, 1000, 2000 and 4000 Hz (AC) and 1000 Hz (BC) were generated by the CTE module of HEARLab and presented to adult participants only using EAR tone 3A insert earphones and a B71 bone conductor.

The speech stimuli of /m/, /t/ and /g/ (see Part A 3.1.1.1 for complete description) were generated by the ACA module of HEARLab and presented in the free-field to infant participants only. The three stimuli were interleaved automatically in blocks of 25 presentations per stimulus.

1.3 Procedure

CAEP recordings were made from Neuroscan™ and HEARLab using a purpose-built triggering device that enabled simultaneous recording from both devices. Electrode sites were prepared before attaching single Ambu ‘Blue Sensor’™ self adhesive, disposable ECG electrodes to Cz (active), M1(reference) and Fz (ground). As each recording system required separate electrode leads a purpose-built ‘double adaptor’ was used to enable the connection from a single disposable electrode at each site to two electrodes (i.e., one for each recording system). The weight of the electrode adaptor and leads was such that an elastic bandage or a headband was used, across the subject’s vertex and forehead, to give more stability to the electrode connections. Further details regarding electrode application and testing techniques in an infant population are shown in Appendix 1.

Electrode impedance was checked, via each test system, and if necessary the preparation was repeated to achieve a recommended impedance under 5 k Ω at all electrode sites using both devices (Roger & Thornton, 2007).

Adults

Cortical testing was performed on two separate occasions with one ear only under stimulation using EAR tone 3A insert earphones. At the first visit and prior to cortical testing participants were assessed with pure tone audiometry (PTA) using standard Hughson-Westlake plateau techniques and contralateral masking to the NTE as appropriate. AC thresholds in 5 dB steps were determined for both ears using TDH-39 headphones, and also using EAR tone 3A insert phones for 500, 1000, 2000 and 4000 Hz. Thresholds were then determined in 2 dB steps using the insert earphones but only in the ear selected for cortical testing. Unmasked BC thresholds were obtained at 1000 Hz in 5 dB steps, and then repeated in 2 dB steps at 1000 Hz. Prior to the cortical testing at the second visit, PTA thresholds were re-established for 1000 and 4000 Hz to ensure that hearing thresholds had not changed since the previous visit.

Otoscopy was performed on both occasions to ensure that the ear canals were clear and there were no contraindications to the insertion of EAR foam tips. Tympanometry was also performed using a Grason-Stadler GSI-33 Middle Ear Analyzer to ensure that the middle ear pressure and tympanic membrane compliance were stable across visits.

During cortical testing adults were seated in a high-backed armchair which could be reclined as desired. Participants were awake and alert during testing and watched a captioned DVD or read a book of their choice. The examiner monitored the participant's state constantly via a video camera to ensure that they remained alert.

All participants had symmetrical hearing thresholds and therefore the left ear was routinely selected for stimulation during cortical testing. Cortical responses were recorded using five presentation levels for each nominated frequency and transducer, making a total of 25 averaged responses per participant. For the purposes of this report however, responses to a single stimulation level will be analysed.

The presentation levels were -10, 0, 10, 20 and 30 dB relative to the appropriate pure tone threshold (measured in 2 dB steps but rounded to the nearest 5 dB). Testing at each presentation level continued until 100 accepted epochs had been recorded in both systems. For some normal hearing participants the lowest sensation levels could not be assessed as the required presentation level was less than the lower limits of HEARLab's range. The order of presentation for air conducted frequencies was balanced across participants with BC testing for 1000 Hz performed last of all. For half the participants, the presentation levels for each stimulus frequency were delivered in an ascending order of sensation level and for the remaining participants the order was reversed to reduce the potential effects of habituation. Contralateral masking was applied to the NTE at 30 dB less than the stimulus presentation level in the TE throughout testing.

Infants

Infants were assessed on one occasion only. Otoscopy and tympanometry were routinely attempted but testing was abandoned if it caused distress. Cortical testing was performed while participants were awake and seated either on their parent/caregiver's lap, or at their side in a large armchair. A trained observer was with the participant and parent in the test booth, to provide distraction, monitor and control the ambient noise level, and to ensure that the electrodes remained well placed during the test. Testing was paused or discontinued if the participant was overly active, too vocal, distressed, obviously drowsy, or losing alertness. Infants were distracted using quiet toys, books, and other visual stimulation such as DVDs without sound.

Test stimuli were presented by a loudspeaker positioned at 0° azimuth and 1.8 m from the test position. Testing continued until a minimum of 90 accepted epochs had been achieved in HEARLab per stimulus at 65 dB SPL. If the infant's state permitted, a second run at 65 dB SPL was attempted. HEARLab automatically separated and displayed the averaged cortical responses to the three interleaved stimuli. To enable the same separation in Neuroscan™ however, recordings were made in continuous streaming with a sound level meter (SLM) positioned near the loudspeaker and connected to the second active channel of the Neuroscan™ head-box. The SLM recordings provided a marker to the change in stimulus but the examiner also manually noted the time of onset for each stimulus block. After acquisition, these

continuous recordings were processed to form epoched files that were a) exported to MATLAB for analysis and b) transformed by the application of baseline correction (see Part A 3.1.3.3), artifact rejection ($\pm 150 \mu\text{V}$), filtering (30 Hz low pass zero-phase filter) and averaging to form a final display averaged response for each stimulus condition.

1.4 Response detection and differentiation

Subjective peak detection

For adult participants, the amplitude and latency of N1 and P2 were marked and recorded for responses at 30 dB SL only, as the clearest morphology was expected at this level. For infant participants, the amplitude and latency of the first positive peak (known as P1) only was recorded for all responses to stimulation at 65 dB SPL. Where peaks could not be unambiguously located, they were identified based on the presence of local maxima (or minima) that preceded and followed the negativity (or positivity). Visual lines of “best-fit” were applied either side of the peak, beginning with the local maxima (or minima) and ending at the nearest local minima (or maxima) within the broad peak region. The point of intersection of the two lines indicated the latency and amplitude mark. As a common point of reference for marking the amplitude of the peak or trough, a second amplitude reading was taken at -100 ms, and the difference between these two measures was reported as the amplitude of the peak. Latency and amplitude readings across the two systems were recorded by different examiners using the peak detection process described.

Objective response detection and differentiation

While HEARLab automatically displayed its difference and detection p values, off-line processing was required for the Neuroscan™ data. In brief, all epoched Neuroscan™ files were exported to MATLAB for analysis. After baseline correction, artifact rejection and filtering (as described under 1.3), the analysis period was determined by age, using the same criteria that was applied in HEARLab (see Part A 3.1.4). Sampling points within the analysis period of each epoch were reduced to nine variables to form a “response” condition which was either analysed using Hotelling’s

T^2 (detection p values) or using MANOVA (difference p values) (see Part A 3.1.4 for details).

2. Results

Infant and adult data were analysed separately due to the unique characteristics of the infant and adult cortical response. Multiple repeated measure analysis of variance (ANOVA) were performed to examine main effects and interactions for the selected test parameters with a p value of < 0.05 used as the point of significance. All adults were assessed on two separate occasions and therefore “visit” was incorporated as a factor in the analysis of adult data. A repeat assessment during the single visit was attempted for all infants but it was only successful in five out of six infants. Of these five cases, wave morphology was poor in response to one or more stimuli in two participants as the children grew tired and restless quickly. As a result, only the first “run” was used in the analysis of infant latency and amplitude data.

2.1 Latency

Infant data

Table 1 shows the mean and standard deviation (SD) for P1 latencies in response to the three selected speech stimuli. When results to all three speech stimuli are combined, the latency of P1 is slightly longer using Neuroscan TM at 168.2 ms (SD 28.3) than HEARLab at 164.7 ms (SD 27.8).

Table 1: P1 latencies (ms) in response to /m/, /t/, /g/ stimuli (N = 6)

	Neuroscan TM	HEARlab
	Mean (SD)	Mean (SD)
/m/	181.4 (36.4)	176 (38.7)
/t/	163.6 (18.4)	164.2 (16.3)
/g/	159.7 (27.1)	153.8 (23.8)

A repeated measure factorial ANOVA was performed to examine HEARLab and Neuroscan TM -generated responses for differences in the latency of P1. The factors

entered to the analysis were stimulus (i.e., /m/, /t/, /g/) and device (i.e., Neuroscan™ and HEARLab). Results showed no significant difference in the latency of P1 across the devices [$F(1,5) = 2.63$; $p = 0.17$]. Similarly, there were no significant differences in latency across stimuli [$F(2,10) = 4.74$; $p = 0.08$], and no interaction between stimuli and device [$F(2,10) = 1.74$; $p = 0.23$].

Adult data

The adult latency results were confined to N1 only, as P2 was poorly formed in many cases making clear peak identification very difficult.

Table 2 shows the mean and standard deviation for N1 latencies in response to four stimuli presented by AC (i.e., AC 500 Hz, AC 1000 Hz, AC 2000 Hz, AC 4000 Hz) and a fifth stimulus presented by BC (i.e., BC 1000 Hz). Some incomplete data sets exist due to calibration errors ($N = 2$) or test-time constraints ($N = 2$). The latency of N1 across all stimuli was consistently longer with Neuroscan™ than HEARLab. At the first visit, the mean N1 latency was 100 ms (SD 9.9) using Neuroscan™ and 93.5 (SD 10) using HEARLab, while at the second visit the mean N1 latency was 101.8 ms (SD 10) using Neuroscan™ and 88.8 (SD 10.5) using HEARLab.

A repeated measures factorial ANOVA was performed to examine responses for differences in the latency of N1. The factors entered to the analysis were stimulus (i.e., AC 500 Hz, AC 1000 Hz, AC 2000 Hz, AC 4000 Hz, BC 1000 Hz), device and visit.

Results showed a significant difference between devices in the latency of N1 [$F(1,7) = 172.34$; $p < 0.001$]. There were no significant effects of stimulus [$F(4,28) = 2.53$; $p = 0.063$], visit [$F(1,7) = 3.65$; $p = 0.10$], and no significant interactions between any of the factors.

Table 2: N1 latencies (ms) in response to five stimuli

	Visit 1		Visit 2	
	Mean (SD)	N	Mean (SD)	N
<i>Neuroscan™</i>				
AC 500 Hz	98.7 (11.3)	12	107.6 (10.1)	11
AC 1000 Hz	97.6 (7)	12	100.8 (8.9)	12
AC 2000 Hz	103.6 (8.3)	12	97.7 (10.7)	12
AC 4000 Hz	105.8 (12.1)	12	105.5 (10.7)	12
BC 1000 Hz	93.5 (4.8)	10	97 (5)	10
<i>HEARLab</i>				
AC 500 Hz	90.8 (12.7)	12	99.9 (9.4)	12
AC 1000 Hz	91.9 (8.5)	12	91.7 (6.5)	11
AC 2000 Hz	98.4 (9.3)	12	91.7 (14.7)	12
AC 4000 Hz	97.6 (9.2)	11	98.4 (11)	12
BC 1000 Hz	88 (5.7)	10	89.7 (6)	10

2.2 Amplitude

Infant data

Table 3 shows the mean and standard deviation for P1 amplitude in response to the three selected speech stimuli.

Table 3: P1 amplitude (μV) in response to /m/, /t/, /g/ stimuli (N = 6)

	Neuroscan TM	HEARlab
	Mean (SD)	Mean (SD)
/m/	9.9 (4.5)	8.9 (5.4)
/t/	10.7 (5.8)	10.3 (4.4)
/g/	10.9 (5.9)	11.4 (5.8)

There appeared to be no consistent difference in the amplitude of P1 across devices. When results were combined for all three stimuli, the mean amplitude of P1 was 10.5 μV (SD 5.1) using Neuroscan TM and 10.2 μV (SD 5) using HEARLab. A repeated measure factorial ANOVA was performed to examine HEARLab and Neuroscan TM - generated responses for differences in the amplitude of P1. The factors entered to the analysis were stimulus (i.e., /m/, /t/, /g/) and device (i.e., Neuroscan TM and HEARLab).

Results showed no significant difference in the amplitude of P1 across the devices [$F(1,5) = 0.19$; $p = 0.68$]. Similarly, there was no effect of stimulus [$F(2,10) = 0.26$; $p = 0.78$] and no significant interaction between stimuli and device [$F(2,10) = 0.46$; $p = 0.64$].

Adult data

Amplitude comparisons will be confined to N1 only as P2 was poorly formed in many instances.

Table 4 shows the mean and standard deviation for N1 amplitude in response to four stimuli presented by AC (i.e., AC 500 Hz, AC 1000 Hz, AC 2000 Hz, AC 4000 Hz) and a fifth stimulus presented by BC (i.e., BC 1000 Hz). Some incomplete data sets exist due to calibration errors (N = 2) or test-time constraints (N = 2).

Table 4: N1 amplitudes (μV) in response to five stimuli

	Visit 1		Visit 2	
	Mean (SD)	N	Mean (SD)	N
<i>NeuroscanTM</i>				
AC 500 Hz	-4.7 (3.1)	12	-5.2 (3.8)	11
AC 1000 Hz	-3.6 (2.2)	12	-4 (2.8)	12
AC 2000 Hz	-4.7 (2.4)	12	-4.5 (2.7)	12
AC 4000 Hz	-4.8 (1.9)	12	-3.6 (2.1)	12
BC 1000 Hz	-5.0 (2.1)	10	-4 (2.8)	10
<i>HEARLab</i>				
AC 500 Hz	-4.9 (3.2)	12	-5.2 (4.5)	12
AC 1000 Hz	-4.1 (2.9)	12	-5 (3.3)	11
AC 2000 Hz	-4.6 (1.7)	12	-4.7 (3.7)	12
AC 4000 Hz	-5.2 (1.9)	11	-4.2 (2.4)	12
BC 1000 Hz	-4.5 (2.3)	10	-4.7 (2.9)	10

The amplitude of N1 responses appears to be similar across visits and systems. When results from visit 1 were combined across stimuli, the mean amplitude was $-4.38 \mu\text{V}$ (SD 2.35) using NeuroscanTM and $-4.65 \mu\text{V}$ (SD 2.46) using HEARLab. For visit 2, the mean amplitude was $-4.23 \mu\text{V}$ (SD 2.79) using NeuroscanTM, and $-4.84 \mu\text{V}$ (SD 3.38) using HEARLab. A repeated measure factorial ANOVA was performed to examine HEARLab and NeuroscanTM-generated responses for differences in the amplitude of N1. The factors entered to the analysis were stimulus (i.e., AC 500 Hz, AC 1000 Hz, AC 2000 Hz, AC 4000 Hz, BC 1000 Hz), device and visit.

Results showed no significant difference between devices in the amplitude of N1 [$F(1,7) = 5.58$; $p = 0.05$] although the outcome was borderline. There were no significant differences in amplitude between stimuli [$F(4,28) = 1.72$; $p = 0.17$], and no significant differences across visits [$F(1,7) = 0.05$; $p = 0.83$]. Similarly, there were no significant interactions between factors, although the stimulus and system interaction was borderline [$F(4,28) = 2.72$; $p = 0.05$].

2.3 Detection P values

All detection P values were calculated using the Hotelling's T^2 statistic and converted to standardized values (z scores) for analysis. The second assessment which was performed in 5 out of 6 infant cases will be included in this analysis.

Infant data

Figure 1 shows the relationship between the z scores for both systems with Neuroscan TM values on the Y-axis and HEARLab on the X-axis. A trend towards higher z scores (i.e., less significant P values) for the second assessment can be seen. With z scores below -3.09 approximately (equivalent to p values < 0.001) there is an apparent shift away from the line of equality with the Neuroscan TM data-generated z scores being lower than those generated by HEARLab. Correlation results for the first run only were high in response to all three stimuli but they were significant in response to /m/ and /g/ only (/m/: $r = 0.94$; $n=6$; $p = 0.005$, /t/: $r = 0.74$; $n = 6$; $p = 0.09$, /g/: $r = 0.96$; $n = 6$; $p = 0.002$). Correlation results for the second run were however high and significant for all stimuli (/m/: $r = 0.95$; $n= 5$; $p = 0.01$, /t/: $r = 0.98$; $n = 5$; $p = 0.004$, /g/: $r = 0.94$; $n = 5$; $p = 0.02$).

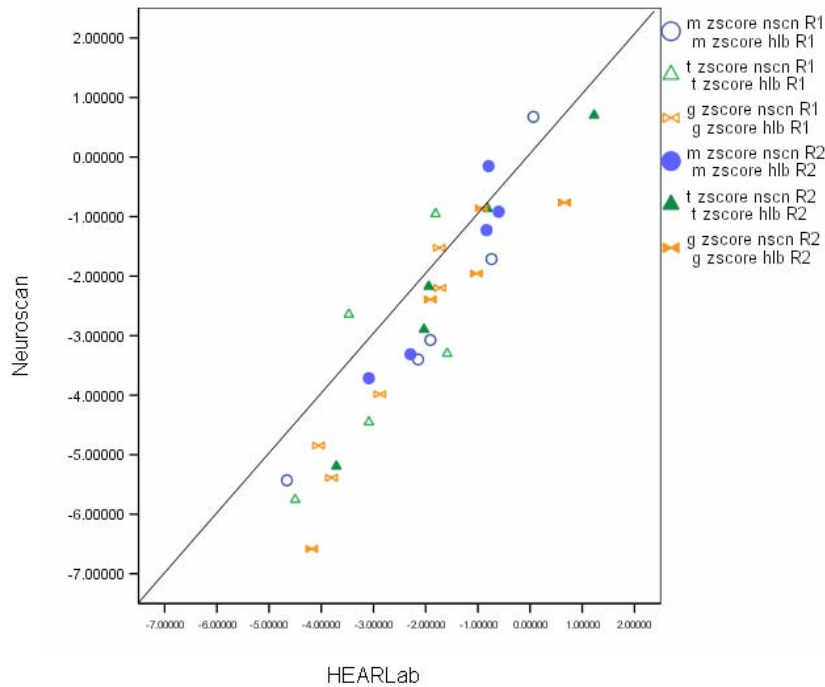


Figure 1: The relationship between z scores (derived from detection P values) for cortical responses generated to three speech stimuli is shown for assessment run one (R1) and two (R2). The z scores for HEARLab-generated responses are shown on the X-axis while the z scores for NeuroscanTM-generated responses are shown on the Y-axis.

A repeated measures factorial ANOVA with stimulus, device and run entered to the analysis showed a significant difference in the z scores between devices [$F(1,4) = 8.39$; $p = 0.04$] and a significant difference in the z scores between runs [$F(1,4) = 11.49$; $p = 0.03$]. There were however no significant differences in the z scores between stimuli [$F(2,8) = 0.66$; $p = 0.54$] and no significant interactions between factors.

Adult data

Figure 2 shows the relationship between the z scores for both systems with NeuroscanTM values on the Y-axis and HEARLab in the X-axis. With z scores below -3.09

approximately (equivalent to p values < 0.001) there is an apparent shift away from the line of equality with the Neuroscan™ data-generated z scores being lower than those generated by HEARLab. It also appears that z scores for the detection p value at 4k Hz are consistently greater than -2 across both systems and generally higher than z scores to other frequencies.

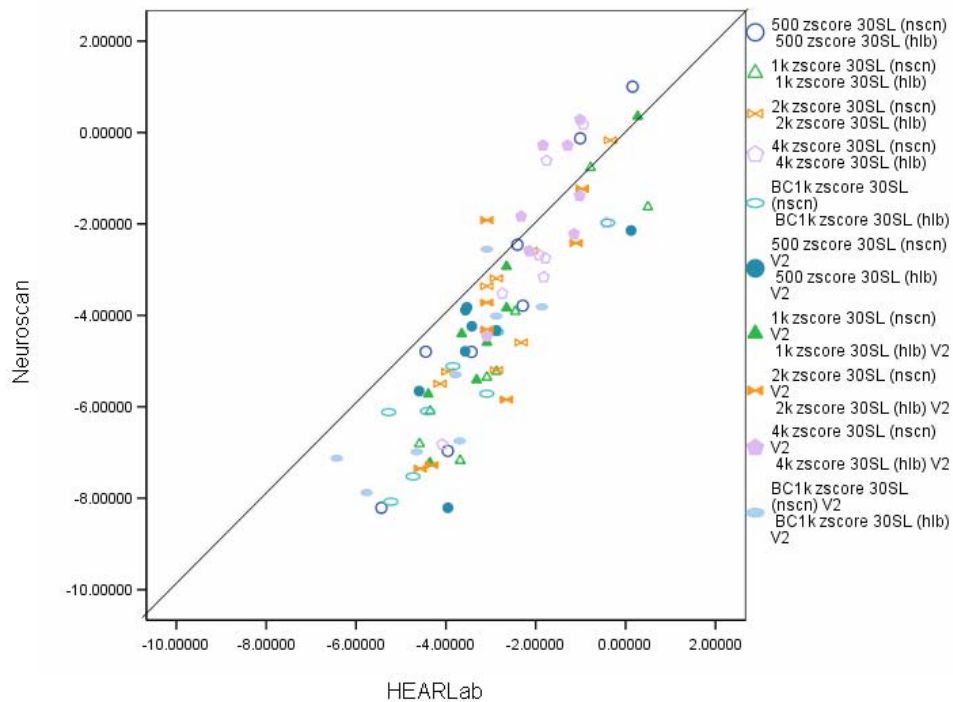


Figure 2: The relationship between z scores (derived from detection p values) for cortical responses generated to five different tonal stimuli is shown for visit one and two (V2).

Table 5 shows the correlations between device - generated z scores for five stimuli and both visits. Correlation results were moderately high or better across both visits and ranged from 0.75 to 0.94, except for AC 4000 Hz at the second visit where the correlation was poor. When results were averaged across all stimuli, the z scores were lower (i.e., the P values were more significant) and had greater variance using Neuroscan™. At the first visit the mean z score was -3.95 (SD 2.19) using Neuroscan

TMand -2.65 (SD 1.47) using HEARLab. At the second visit, the mean z score was -3.92 (SD 2.18) using Neuroscan TM and -2.72 (SD 1.47) using HEARlab.

Table 5: Correlations (r) between Neuroscan TM and HEARLab-generated z scores using five stimuli

	Visit 1	N	Visit 2	N
AC 500 Hz	0.91**	12	0.75**	11
AC 1000 Hz	0.82**	12	0.94**	12
AC 2000 Hz	0.84**	12	0.83**	12
AC 4000 Hz	0.78**	12	0.47*	12
BC 1000 Hz	0.87**	10	0.82**	10

** p < 0.01

* p > 0.05

A repeated measures factorial ANOVA with stimulus, device and visit entered to the analysis showed a significant difference in the z scores between devices [$F(1,7) = 19.41$; $p = 0.003$] and a significant difference in the z scores between stimuli [$F(4,28) = 13.33$; $p < 0.001$]. Bonferroni pair-wise comparisons showed a significant difference between AC 500 and 4000 Hz ($p = 0.02$), AC 500 and BC 1000 Hz ($p = 0.04$), AC 1000 and BC 1000 Hz ($p = 0.03$), and AC 4000 and BC 1000 Hz ($p = 0.001$). When two further repeated measure factorial ANOVAs were performed with stimulus and device entered to the analysis only, a significant difference in the z scores between stimuli was observed again (visit 1 [$F(4,36) = 7.18$; $p < 0.001$]; visit 2 [$F(4,32) = 10.29$; $p < 0.001$]) however the Bonferroni pair-wise comparisons showed a different pattern between visits. For visit 1, there was a significant difference between AC 4000 and BC 1000 Hz only ($p = 0.01$) but for visit 2 there were significant differences between AC 500 and AC 4000 Hz ($p = 0.01$), AC 2000 and AC 4000 Hz ($p = 0.02$) and AC 4000 and BC 1000 Hz ($p < 0.001$).

There was also a significant interaction between stimulus and device [$F(4,28) = 5.51$; $p = 0.002$] but no significant difference in z scores between visits [$F(1,7) = 0.054$; $p =$

0.823], and no other significant interactions between factors. The significant interaction between stimulus and device is shown in Figure 3. The lower z scores (i.e., more significant P values) for Neuroscan™-generated responses are evident for all stimuli except 4000 Hz where the z scores are similar using both systems.

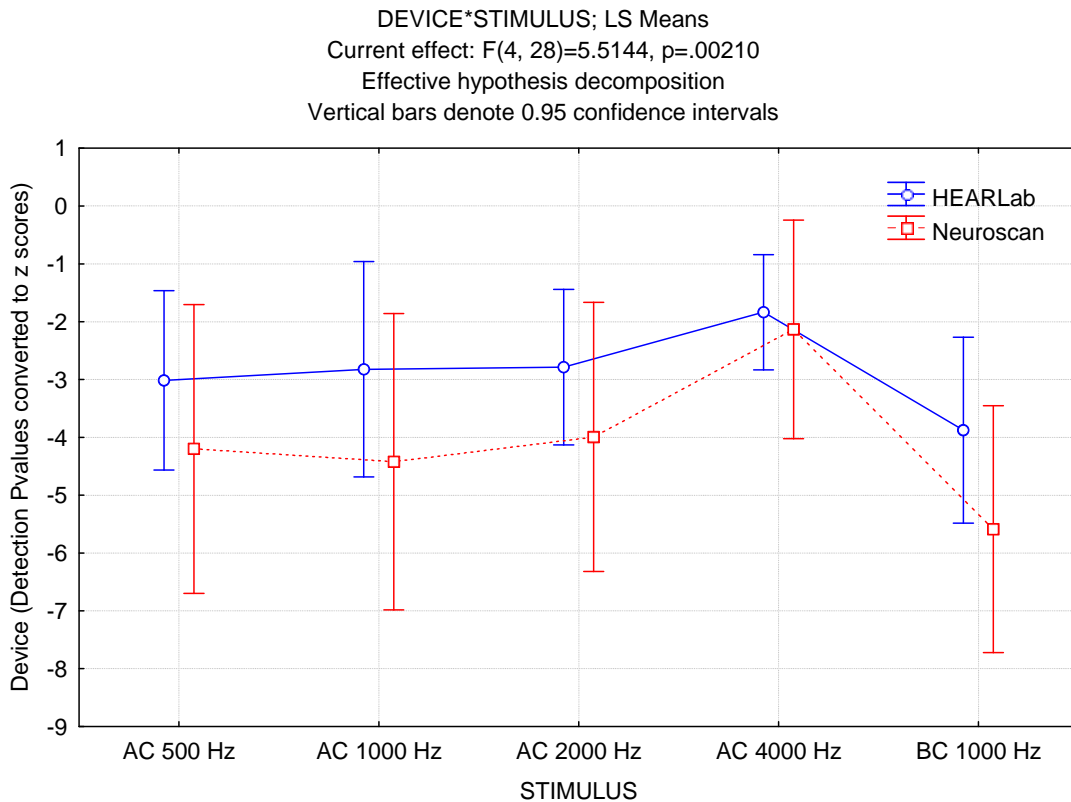


Figure 3: A significant interaction between the z scores (derived from detection P values) for cortical responses generated to five stimuli and two devices is shown. The z scores derived for HEARLab-generated responses are higher than those for Neuroscan™-generated responses except for AC 4000 Hz where the values are similar.

Detection P values, from Visit 1 and 2, were used to divide the HEARLab - and Neuroscan- generated responses into categories representing the presence/absence of a cortical (based on $p<0.05$) for stimuli presented at 30 dB SL. Possible categories were a) present using both systems, b) absent using both systems, c) present using Neuroscan™ but absent using HEARLab, or d) present using HEARLab but absent

using Neuroscan. Table 6 shows that the two systems were in agreement for 83% of responses but disagreed for the remainder. For this subset of 20 cases, HEARLab reported that the response was absent on 17 occasions while NeuroscanTM reported that the responses were absent on three occasions.

Table 6: A comparison of cortical detection rate across systems for stimuli presented at 30 dB SL, during Visit 1 and 2.

Stimulus	Systems agree		Systems disagree		N
	Both pres*	Both abs**	Scan pres/ HLab abs	NScan abs/ HLab pres	
AC 500Hz	19	2	2		23
AC 1000Hz	18	4	1	1	24
AC 2000Hz	17	3	4		24
AC 4000Hz	11	4	7	2	24
BC 1000Hz	17	0	3		20
Total	82	13	17	3	115

* present

** absent

2.4 Difference P values

All difference P values were calculated using MANOVA and converted to standardised values (z scores) for analysis. The difference p values were calculated using /m/, /t/, and /g/ stimulus-generated responses from infants only. The second assessment which was performed in 5 out of 6 infant cases will be included in this analysis.

Infant data

Figure 4 shows the relationship between the z scores for both systems with NeuroscanTM values on the Y-axis and HEARLab on the X-axis. There are no apparent differences in outcomes between the first and second run. Correlation results for the

first run only were high and significant for the /gm/ and /tm/ pairs whereas the /gt/ pair had a high correlation but was not significant (/gm/: $r = 0.97$; $n=6$; $p = 0.002$, /tm/: $r = 0.90$; $n = 6$; $p = 0.015$, /gt/: $r = 0.78$; $n = 6$; $p = 0.07$). Correlation results for the second run were also high and significant for the /gm/ and /gt/ pairs whereas the correlation for /tm/ was poor and non significant (/gm/: $r^s = 0.90$; $n = 5$; $p = 0.04$, /tm/: $r = 0.33$; $n = 5$; $p = 0.59$, /gt/: $r^s = 0.90$; $n = 5$; $p = 0.04$).

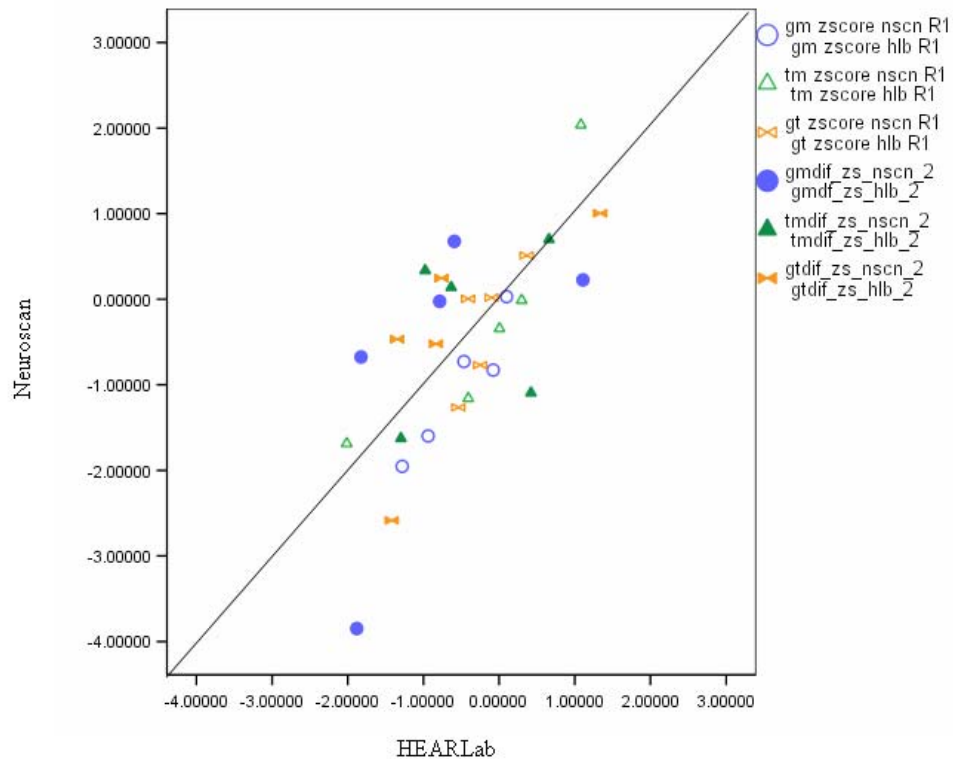


Figure 4: The relationship between z scores (derived from difference P values) for cortical responses generated to speech stimuli pairs is shown for assessment run one (R1) and two (R2). The z scores for HEARlab-generated responses are shown on the X-axis while z scores for Neuroscan™ - generated responses are shown on the Y-axis.

A repeated measures factorial ANOVA with stimulus, device and run entered to the analysis showed no significant difference in the z scores between devices [$F(1,4) =$

0.09; $p = 0.79$], nor between runs [$F(1,4) = 0.19$; $p = 0.69$] and no significant difference in the z scores between stimulus pairs [$F(2,8) = 0.86$; $p = 0.46$]. Similarly, there were no significant interactions between any of these factors.

Difference P values, for Run 1 and 2, were used to divide the HEARLab – and Neuroscan TM – generated responses into categories representing the presence/absence of a difference in the cortical response to paired stimuli. Possible categories were a) difference detected using both systems or no difference detected using both systems (i.e., system agreement), b) difference detected using Neuroscan TM but no difference detected using HEARLab, or c) difference detected using HEARLab but no difference detected using Neuroscan TM. Table 7 shows that the systems were in agreement for 91% of paired responses.

Table 7: A comparison of the cortical response differences for all stimulus pairs across systems and both runs.

Stimulus	Systems agree	Systems disagree		N
		NScan pos/ HLab neg	NScan neg/ HLab pos	
/g/ vs /t/	10	1	0	11
/g/ vs /m/	9	1	1	11
/t/ vs /m/	11	0	0	11
Total	30	2	1	33

2.5 Reject rate

The proportion of rejected responses was compared across devices using both infant and adult data sets. There are clear differences in the overall reject rate between both populations with adults having much lower rates than infants. The reject rate in infants when results were combined for all three stimuli was 0.11 (SD 0.1) for Neuroscan TM - generated responses and 0.16 (SD 0.10) for HEARLab-generated responses in infants. When adult data was combined across visits and stimuli, the overall reject rate was

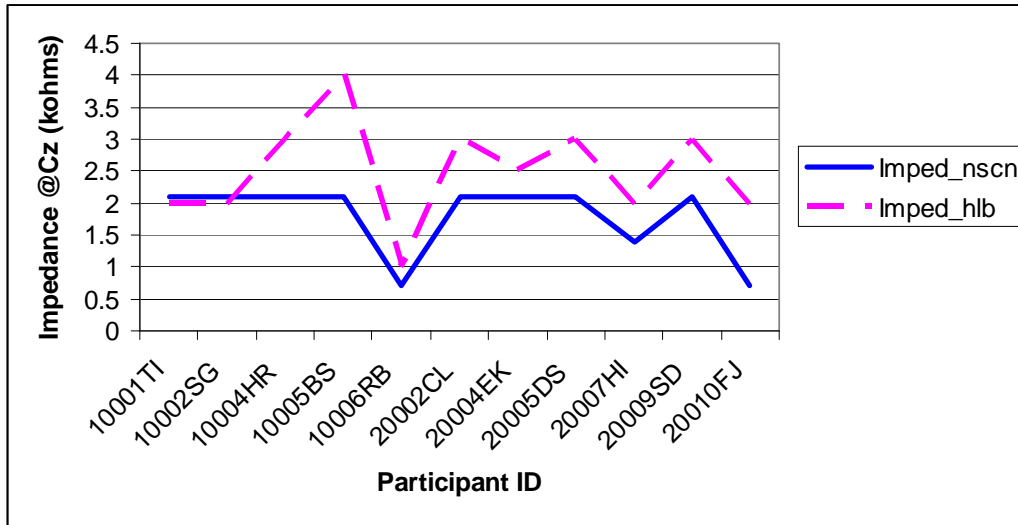
very low at 0.0005 (SD 0.003) in Neuroscan™ and 0.01 (SD 0.02) in HEARLab. Further result analysis will therefore be confined to the infant data.

A repeated measures factorial ANOVA with stimulus, device and run entered to the analysis showed no significant difference in the rate of rejection between devices [$F(1,4) = 6.83$; $p = 0.06$], no significant differences across stimuli [$F(2,8) = 1.13$; $p = 0.37$] and no significant differences across runs [$F(1,4) = 0.08$; $p = 0.79$]. Similarly, there were no significant interactions between factors.

2.6 Impedance

While HEARLab provides an impedance check of the active and reference sites, the default impedance measurement in Neuroscan™ is taken at the active site only in single channel recordings. As a result, the impedance at Cz was compared across devices. It is feasible in Neuroscan™ but not in HEARLab to record a numeric for impedance as well as interpolate a value from a coloured scale. As a result the impedance estimates were based on readings from the coloured scale in each device. As a low impedance was required before testing commenced, there was little variability across participants and hence raw results are shown in Figure 5. The difference in the impedance at the Cz site is consistently lower in Neuroscan™ than in HEARLab.

A



B

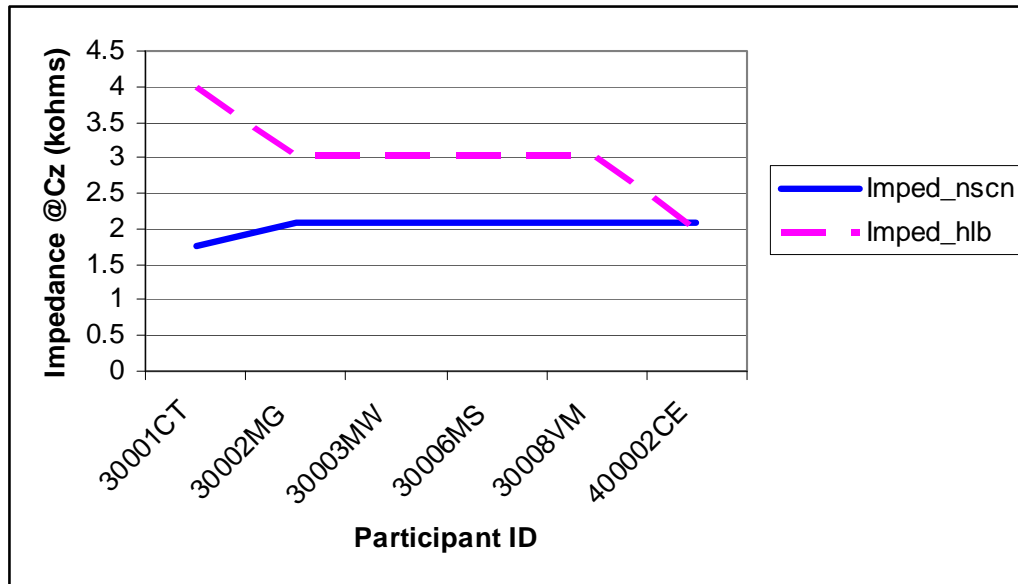


Figure 5: The impedance recorded at the Cz site for HEARLab and Neuroscan™ devices is shown. Results show higher impedance from HEARLab than Neuroscan™ for adult participants (A) and for infant participants in (B).

2.7 Residual noise (“RN”)

The RMS value of the residual noise that is superimposed on the averaged accepted epochs is calculated concurrent with the detection p value for each stimulus-generated response (see Part A: 3.1.4.3).

Infant

The mean (SD) of “RN” is shown in Table 8. The results appear to be similar across assessments but the mean values appear to be slightly lower from the HEARLab than from the Neuroscan™ - generated responses.

Table 8: Mean (μ V) of “RN” by stimulus and device

	Run 1 (N = 6) Mean (SD)	Run 2 (N = 5) Mean (SD)
<i>Neuroscan™</i>		
/m/	3.8 (0.5)	3.6 (0.7)
/t/	3.8 (0.6)	3.7 (0.8)
/g/	3.7 (0.6)	3.8 (0.5)
<i>HEARLab</i>		
/m/	3.6 (0.5)	3.5 (0.5)
/t/	3.6 (0.4)	3.5 (0.5)
/g/	3.5 (0.5)	3.6 (0.4)

A repeated measures factorial ANOVA with stimulus, device and run entered to the analysis showed no significant difference in “RN” between devices [$F(1,4) = 2.89$; $p = 0.17$], nor stimuli [$F(2,8) = 0.65$; $p = 0.55$] and no significant difference between assessments [$F(1,4) = 0.01$; $p = 0.91$]. Similarly, there were no significant interactions between factors except for stimulus and device [$F(2,8) = 4.78$; $p = 0.04$]. This interaction is illustrated in Figure 6.

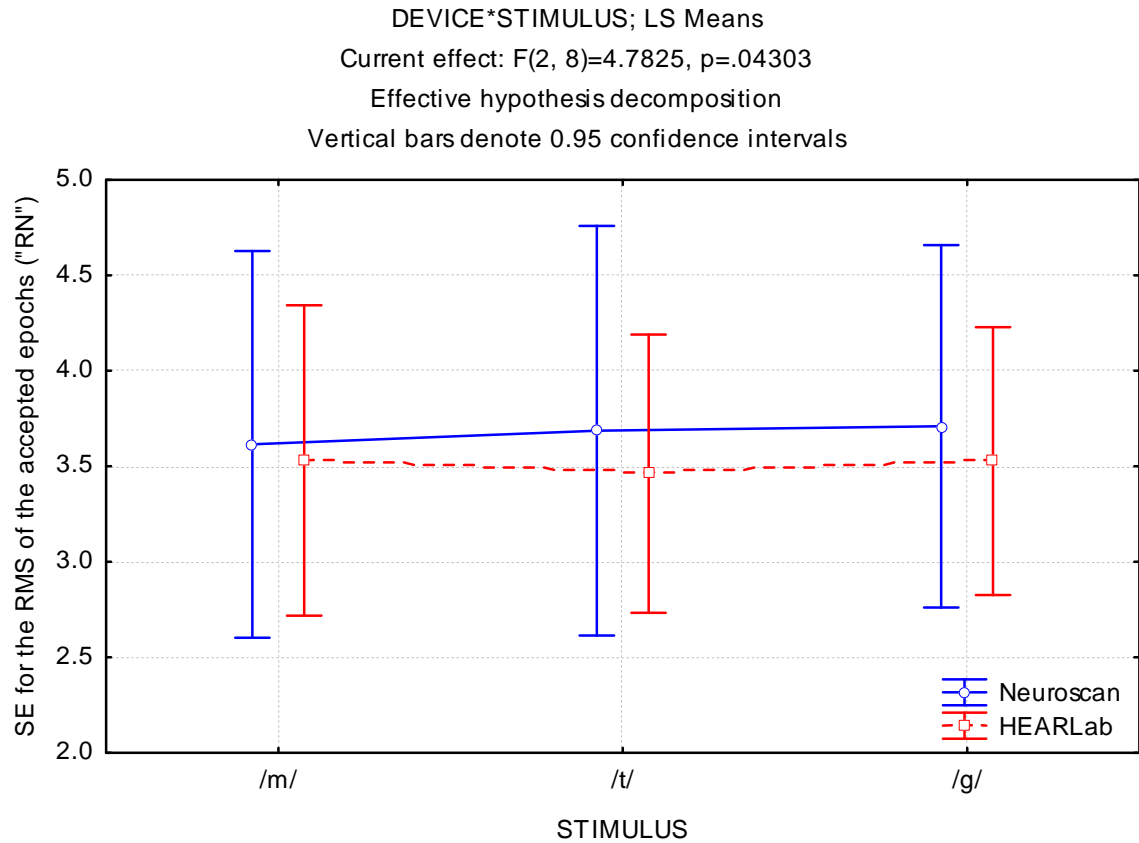


Figure 6: A significant interaction between the SE for the dB RMS of the accepted epochs generated to the three speech stimuli and two devices is shown.

3. Discussion

The infant obligatory CAEP in response to speech stimuli consists of one or more positive peaks with the most prominent and first peak occurring at 200 to 300 ms in newborns when recorded at the midline (Sharma et al., 2002; Stapells et al., 1991). In this report, the amplitude and latency results will therefore be confined to the first positive peak known as P1. The adult obligatory cortical response consists mainly of N1 and P2 but response amplitude, latency and morphology are known to vary substantially across individuals (Hyde, 1997). In our study it was common for many responses to show a poorly formed P2. This is consistent with earlier research where P2 has been reported as “little more than a recovery at the baseline” (Hyde, 1997). This made the application of the peak detection criteria, as described in this study, very difficult and resulted in many instances of missing data for P2 amplitude and latency. As a result, the amplitude and latency results are confined to the most stable component of the adult result which is N1 (Beagley & Knight, 1967).

3.1 Latency

Our results show that the latency of the infant P1 response was not significantly different across devices and when the P1 latency of the infant responses were plotted in HEARLab, they all fell within the age-appropriate latency band which is provided in the device as a result of earlier work (Golding et al., 2007). When the adult data was examined there was however a significant difference in the latency of the adult N1 response between devices. Although the same peak-latency identification technique was used for both the infant and adult data sets, the adult responses were less noisy (see the results for artifact reject rates) which reduced variance and made adult peak detection more precise. It is likely that there is a genuine difference in the latency measure between devices with the N1 latency being, on average, 7 – 13 ms shorter in HEARLab than Neuroscan™. This difference is however likely to be of little consequence for the unit’s clinical purpose of auditory threshold estimation particularly as the N1 latency range in normal hearing adults has been reported to be wide at 80 – 120 ms (Davis, 1965).

Of secondary interest was the finding that there were no significant differences in the latency of P1 (infant results) or N1 (adult results) with changes in stimuli. Some previous research has suggested that stimuli with a lower frequency emphasis should generate cortical responses with longer latencies than high frequency stimuli in normal hearing adults (Jacobson, Lombardi, Gibbens, Ahmad, & Newman, 1992) but others have reported little or no effect of latency change with changes to the stimulus frequency (Hyde, 1997; Zerlin & Naunton, 1974; Agung, Purdy, McMahon, & Newall, 2006). There appears then to be conflicting evidence for a simple effect of stimulus frequency on latency change in CAEPs.

3.2 Amplitude

Our results showed no significant difference between devices in the amplitude of the infant P1 or adult N1 response. Of secondary interest were the findings that a) there was no difference in amplitude of P1 (infant data) or N1 (adult data) with changes in stimuli, and b) there were no differences in the N1 amplitude across visits.

Previous research has suggested that higher frequency stimuli are likely to generate responses with lower amplitude than those resulting from low frequency stimuli in normal hearing adults (Agung et al., 2006; Jacobson et al., 1992). These findings of amplitude change with stimulus frequency are consistent with fMRI studies of the tonotopic organisation of the cortex where cortical regions that respond to low frequency information are located closer to the surface of the scalp than those regions associated with high frequency stimuli (Yetkin, Roland, Christensen, & Purdy, 2004). It should be recalled however that we assessed infants and adults with normal hearing and hearing impairment, and the application of findings from adults with normal hearing to other populations is fraught with difficulties. In addition, substantial differences in measurement protocols across studies is likely to result in differing outcomes because the obligatory CAEPs is highly susceptible to the interactive effects of perceived changes in the auditory environment (Hyde, 1997).

Amplitude is known to be highly susceptible to the individual's level of alertness (Hyde, 1997; Davis, 1965). In this study, participants were therefore engaged in watching a DVD or reading a book of their choice during each recording session and

monitored constantly for drowsiness (see “1.3 Procedures” for a complete description). Although it is highly unlikely that all participants would be equally drowsy on one visit and equally alert at the next, the stability of N1 amplitude across visits was a positive outcome.

3.3 Detection P values

Our results, from infant and adult data sets, showed a significant difference in the detection P value across devices with more conservative (i.e., higher) P values for HEARLab - than Neuroscan™ - generated responses. Similarly, the rate of reporting absent cortical responses, when responses were expected to be present, was twice as high in HEARLab compared with Neuroscan. Of secondary interest were the findings that a) when the z scores were compared across runs (infant data) or across visits (adult data) there were significant differences observed in the infant data only, b) there was no significant difference in the detection P z scores between the speech stimuli that were used in the infant study but there was a significant difference in the detection P z scores for the tonal stimuli used in the adult study.

The difference in the detection P values across devices is not completely surprising as the implementation of the Hotelling's T^2 statistic is not identical in the two systems. HEARLab performs an *on-line* analysis which is updated regularly as ten pairs of accepted epochs are acquired. Each pair consists of an accepted response to a positive- and negative-polarity stimulus and the paired data is averaged to form a single data set per pair. This pairing may or may not be based on adjacent epochs depending on the rate of rejection within each recording. In contrast, the final set of epoched data that is acquired using Neuroscan™ is processed *off-line* using a MATLAB script and responses to all stimuli are simply treated as single sets for analysis. There is, therefore, potentially double the number of data sets entering the Hotelling's T^2 analysis for Neuroscan – generated responses than is the case for HEARLab –generated responses. The pairing of responses was implemented in HEARLab because physiological differences to the different stimulus polarities have been noted (at least at the level of the brainstem)(Hall, 1992) and pairing is effective in reducing the contaminating effects of stimulus artifact (Hall, 2007; Goldstein et al., 1999). This latter effect is not however critical to recording late potentials (Hall,

2007). It is possible that while there may be some advantage to pairing, the associated reduction in the number of datasets available to Hotelling's T^2 may have had an adverse effect on the outcome.

There are a number of other differences in the way that the two systems produce their detection P value. Firstly, a DC offset was applied to the pre-stimulus region in HEARLab before the application of a number of noise rejection criteria, whereas Neuroscan™ – generated responses were recorded using an AC coupled filter and a single noise rejection criterion of $\pm 150 \mu\text{V}$. Although the noise rejection criteria differences are substantial, the final number of rejected epochs for each visit or assessment was found to be very similar across devices and therefore this system variation is unlikely to explain the difference in detection P values. Secondly, Neuroscan™ -generated responses were recorded using a 0.1 – 30 Hz band -pass filter and another 30 Hz low-pass zero-phase filter was applied to the accepted epochs during off-line processing. HEARLab applies a high-pass filter of 0.3 Hz in the hardware and a 30 Hz low-pass zero-phase filter in the software after the application of the DC offset and noise rejection criteria. These system variations may have caused some subtle differences in the detection P values but are unlikely to explain the system variation in the rate of reporting absent cortical responses. Finally, the method by which P values were converted to z scores may have inadvertently introduced a measurement error. A review of the relevant scatterplots for both the infant and adult data sets (i.e., Figures 1 and 2) shows a shift away from the line of equality towards higher z score values in HEARLab than Neuroscan™ . This trend became evident when z scores decreased below -3.0 (an equivalent of $p < 0.001$). This effect may have occurred because no limit was placed on the number of decimal points for the P value before conversion to the z score and each system had a variable number of decimal points. While this measurement error might account for the system variation when P values were very small, it is of little clinical consequence.

A secondary finding from our study was that significant differences in the detection P value were evident when z scores were compared across runs using the infant data but not when z scores were compared across visits using the adult data. The stability of the response detection across visits was expected using the adult data, because stimuli were presented well above auditory threshold on both occasions. For the infant study,

speech stimuli were presented at a constant 65 dB SPL during both runs. This level was expected to be well above auditory threshold for five out of six infants and therefore clear cortical responses were expected. A review of the scatter-plot for infant data (see Figure 1) shows a higher number of z scores greater than -1.64 ($p > 0.05$) in run 2 than in run 1 suggesting a higher number of absent responses during the second run. This was consistent with observations made by the examiner that the infants grew tired of testing rapidly and hence the likelihood of response detection was reduced.

The other secondary finding was that the detection P z scores were significantly different within the set of tonal stimuli but within the set of speech stimuli. Figures 2 and 3 show a trend towards high z scores (i.e., p values are less likely to be significant) for AC 4000 Hz stimulation, particularly at visit 2, and low z scores (i.e., p values are more likely to be significant) to BC 1000 Hz stimulation. Post hoc comparisons however showed that statistically significant differences in the detection P z scores were confined to a) AC 4000 and BC 1000 Hz only at the first visit, b) AC 500 and AC 4000 Hz, AC 2000 and AC 4000 Hz, AC 4000 and BC 1000 Hz at the second visit. The AC 4000 Hz and BC 1000 Hz pair was therefore the only one to show a consistent difference in the detection P z score for both visits. Given that the presentation level for all stimuli was 30 dB SL, the relatively high z score associated with AC 4000 Hz tends to suggest a poorer SNR for this frequency than any other. This is consistent with previous reports of small N1-P2 amplitude responses for high compared with low frequency stimulation (Agung et al., 2006; Jacobson et al., 1992), and less amplitude growth with increasing intensity using high frequency stimuli (Hyde, 1997). It should also be noted that the correlation between Neuroscan™ - and HEARLab -generated z scores was poor for AC 4000 Hz at visit 2 only whereas all other correlations were significant and at least moderately high. While there may be a case for some electrophysiological advantage for lower frequency stimuli in cortical response detection, the lack of consistency in the correlation of z scores from Visit 1 to Visit 2 for AC 4000 Hz cannot be attributed to this or system differences.

3.4 Difference P values

The difference P values were calculated for the infant responses to the stimulus pairs /m/ vs /t/, /t/ vs /g/, and /g/ vs /m/ only. A MANOVA analysis (see Part A: 3.1.4.2) was applied to the stimulus pairs across both systems using the processes described previously for detection P value calculation.

Unlike the outcomes observed to the detection p values, there was generally good agreement between devices in the categorising of outcomes as “difference present” or “difference absent”. Similarly, there was no significant variation between devices in the difference P z scores. The scatterplot comparison of the relationship between the z scores for both systems (see Figure 4) shows that there were only two cases where the z score was less than -2.0 for one or both systems. This is in contrast to the z scores associated with the detection P values where the number of data points above and below -2.0 is more evenly distributed. The decimal-point rounding dilemma that was experienced with the z score conversion for very small detection P values was therefore not an issue.

Although of secondary interest to this study, it should be noted that the majority of data points for difference P values were greater than -1.64 (i.e., a p value > 0.05) suggesting that there were very few significant differences within the stimulus pairs, and similarly there were no significant differences in the z scores across stimulus pairs (i.e., no pair was more easily differentiated than another). This appears to be inconsistent with earlier reports from this laboratory where responses from 20 infants with normal hearing were evaluated. In that earlier study nearly all infants (i.e., 19 out of 20) had statistically different responses for /g/ vs /m/ and for /t/ vs /m/ but only half had /t/ vs /g/ differences (Purdy et al., 2005; Dillon, 2005). The smaller sample size in the current study and lack of group homogeneity with respect to hearing status may account for some of this discrepancy but further research on the difference P value is needed.

3.5 Reject rates and residual noise

As expected, the number of rejected epochs was higher in infants than in adults but there was no difference in the rejection rate across systems. Similarly, there was no

significant difference in the RMS of the residual noise between systems even though it appeared, from Table 6, that the residual noise was slightly lower in HEARLab than in Neuroscan™. The likelihood for detecting a cortical response given the number of accepted epochs and the level of electrical activity within them was therefore similar across systems.

4. Conclusion and recommendations

The overall aim of this study was to a) examine the suitability of HEARLab's selected test parameters and user-interface for CAEP testing, and b) perform a clinical study in which the accuracy of HEARLab's two modules in recording CAEPs was compared with responses recorded simultaneously from a second device (i.e., Neuroscan™) that is known to produce true electrophysiological outcomes.

The majority of components within HEARLab's device requirements specifications were justifiable and require no further action. There were some minor exceptions as follows:

- f) 3.1.1.1.1 speech sounds,
- g) 3.1.1.1.2 duration of speech sounds,
- h) 3.1.1.2.7 masking,
- i) 3.1.3.1 display acquired signal,
- j) 3.1.3.3 calculate and display average accumulated epochs.

The clinical study showed few system differences of note. One exception was a significant difference in the statistical detection P value measure between systems where HEARLab was found to be more conservative than Neuroscan™ in detecting CAEPs. This manifested itself in two ways. Firstly, a measurement error may have occurred from the conversion of very small P values to z scores but this has little clinical consequence. Secondly, and more importantly, when cortical responses to suprathreshold stimuli were categorised as “present” or “absent”, the lack of response detection was much higher in HEARLab than Neuroscan™. Although there are numerous system differences in how raw data is processed and the Hotelling's T^2 statistic is applied, the majority of differences are unlikely to explain this finding and further work is needed to identify the source.

In the first instance, it is recommended that Neuroscan's epochs are re-processed to form 50 datasets (i.e., epochs are paired and averaged as per HEARLab's procedure) and Hotelling's T^2 is re-applied. It is acknowledged that the exact pairing used in HEARLab will not be replicated but fortunately very few epochs in either system were lost to noise rejection and the approximation should be reasonable. A categorical comparison of the outcomes will then be required and if favourable, the policy of pairing accepted epochs in HEARLab should be reviewed. If this investigation proves unsatisfactory, it is recommended that the ability to retain raw epochs be implemented in HEARLab and some new data be recorded concurrently from NeuroscanTM and HEARLab. Raw epochs from both systems should then be exported to MATLAB for processing using approximations of the current HEARLab processing scheme and the current NeuroscanTM scheme. It is hoped that the system difference in detection P values can be re-created and the source identified using this method.

Reference List

- Abramovic, S. (1990). *Electric response audiometry in clinical practice*. Churchill Livingstone.
- Agung, K., Purdy, S. C., McMahon, C., & Newall, P. (2006). The use of cortical auditory evoked potentials to evaluate encoding of speech sounds in adults. *Journal American Academy Audiology*, 17, 559-572.
- Alain, C., Woods, D. L., & Covarrubias, D. (1997). Activation of duration-sensitive auditory cortical fields in humans. *Electroencephalography and Clinical Neurophysiology*, 104, 531-539.

- Beagley, H. A. & Knight, J. J. (1967). Changes in auditory evoked response with intensity. *Journal of Laryngology and Otology*, 81, 861-873.
- Brown, E., Klein, A. J., & Snyder, K. A. (1999). Hearing aid processed tone pips: Electroacoustic and ABR characteristics. *Journal American Academy Audiology*, 10, 190-197.
- Clemis, J. D., Ballard, W. J., & Killian, M. C. (1986). Clinical use of an insert earphone. *Annals of Otology, Rhinology and Laryngology*, 95, 520-524.
- Cone-Wesson, B. & Wunderlich, J. (2003). Auditory evoked potentials from the cortex: audiology applications. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 11, 372-377.
- Davis, H. (1965). Slow cortical responses evoked by acoustic stimuli. *Acta Otolaryngologica*, 59, 179-185.
- Davis, H. & Zerlin, S. (1965). Acoustic relations of the human vertex potential. *Journal Acoustical Society of America*, 39, 109-116.
- Dillon, H. (2005). Page Ten: So, baby, how does it sound? Cortical assessment of infants with hearing aids. *The Hearing Journal*, 58, 10-17.
- Dillon, H., Golding, M., Purdy, S. C., & Katsch, R. (2006). Automated detection of cortical auditory evoked potentials.[Abstract] *The Australian and New Zealand Journal of Audiology*, 28 (Suppl.), 20.
- Durrant, J.D. & Hyre, R. (1993). Relative effective frequency response in bone versus air conduction stimulation examined via masking. *Audiology*, 32: 175 - 184.

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

Gilley, P., Sharma, A., Dorman, M. F., & Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clinical Neurophysiology*, 116, 648-657.

Golding, M. (2006). [A comparison of new versus old speech stimuli]. Unpublished raw data.

Golding, M., Dillon, H., & Seymour, J. (2006). [A "quality of test" measure for implementation in HEARLab]. Unpublished raw data.

Golding, M., Dillon, H., Seymour, J., & Carter, L. (2007). *The sensitivity of Hotelling's T^2 in the detection of obligatory cortical auditory evoked potentials (CAEPs)* Manuscript in preparation.

Golding, M., Dillon, H., Seymour, J., Purdy, S. C., & Katsch, R. (2006). *Obligatory Cortical Auditory Evoked Potentials (CAEPs) in infants - a five year review*. National Acoustic Laboratories Research & Development Annual Report 2005/2006. Chatswood, NSW, Australia: Australian Hearing, 15 - 19.

Golding, M., Pearce, W., Seymour, J., Cooper, A., Ching, T. Y. C., & Dillon, H. (2007). The relationship between obligatory cortical auditory evoked potentials (CAEPs) and functional measures in young infants. *Journal American Academy Audiology*, 18, 117-125.

Golding, M., Purdy, S. C., Sharma, M., & Dillon, H. (2006). The effect of stimulus duration and interstimulus interval on cortical responses in infants. *Australian and New Zealand Journal of Audiology*, 28, 122-136.

Goldstein, R. & Aldrich, A. M. (1999). *Evoked potential audiometry: Fundamentals and Applications*. Allyn and Bacon.

Gordon, J. S., Phillips, D. S., Helt, W. J., Konrad-Martin, D., & Fausti, S. A. (2005). Evaluation of insert earphones for high-frequency bedside ototoxicity monitoring. *Journal of Rehabilitation, Research & Development*, 42, 353-362.

Hall, J. W. (1992). *Handbook of auditory evoked responses*. Allyn and Bacon.

Hall, J. W. (2007). *New handbook of auditory evoked responses*. Boston: Pearson, Allyn and Bacon.

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

Hyde, M. (1994). Signal processing and analysis. In J.T. Jacobson (Ed.), *Principles and applications in auditory evoked potentials* (1st ed., pp. 47-84). Needham Heights: Allyn and Bacon.

Hyde, M. (1997). The N1 Response and Its Applications. *Audiology and Neuro-otology*, 2, 281-307.

Jacobson, G. P., Lombardi, M. A., Gibbens, N. D., Ahmad, B. K., & Newman, C. W. (1992). The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component N1. *Ear and Hearing*, 13, 300-306.

Jasper, H.H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology*, 10, 371 -375.

Killion, M. C. & Villchur, E. (1989). Comments on "Earphones in audiometry". *Journal Acoustical Society of America*, 85, 1775-1778.

Kurtzberg, D. (1989). Cortical event-related potential assessment of auditory system function. *Seminars in Hearing*, 10, 252-262.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., & Naatanen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Cognitive Neuroscience and Neuropsychology*, 13, 47-51.

McPherson, D. L. (1996). *Late Potentials of the Auditory System*.

National Acoustic Laboratories. (1982). *Tables for determining percentage loss of hearing*. Commonwealth of Australia (reprinted 1999).

Onishi, S. & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked V potentials. *Journal of the acoustical society of America*, 44, 582-591.

Purdy, S. C., Katsch, R., Dillon, H., Storey, L., Sharma, M., & Agung, K. (2005). Aided cortical auditory evoked potentials for hearing instrument evaluation in infants. In R.Seewald (Ed.), *A sound foundation through early amplification* (Warrenville, IL: Phonak).

Purdy, S. C., Katsch, R., Storey, L., Dillon, H., Agung, K., & Sharma, M. (2006). Obligatory cortical auditory evoked potentials to speech and tonal stimuli in infants and adults with normal hearing. Manuscript submitted for publication.

Reid, A., & Thornton, A.R.D. (1983). The effects of contralateral masking upon brainstem electric responses. *British Journal of Audiology*, 17, 155-162.

Roger, A. & Thornton, A.R.D. (2007). Instrumentation and recording parameters. In Burkard, R.F., Don., M., & Eggermont, J.J. *Auditory evoked potentials: Basic principles and clinical application*. Baltimore: Lippincott Williams & Wilkins,

Rothman, H. H., Davis, H., & Hay, I. S. (1970). Slow evoked cortical potentials and temporal features of stimulation. *Electroencephalography and Clinical Neurophysiology*, 29, 225-232.

Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear and Hearing*, 23, 532-539.

Simos, P. G. & Molfese, D. L. (1997). Electrophysiological responses from a temporal order continuum in the newborn infant. *Neuropsychologia*, 35, 89-98.

Stapells, D. R. (2002). Cortical event-related potentials to auditory stimuli. In J.Katz (Ed.), *Handbook of Clinical Audiology* (pp. 378-406). Philadelphia: Lippincott Williams and Wilkins.

Stapells, D. R. & Kurtzberg, D. (1991). Evoked potential assessment of auditory system integrity in infants. *Clinics in Perinatology*, 18, 497-518.

Yetkin, F. Z., Roland, P. S., Christensen, W. F., & Purdy, P. D. (2004). Silent functional magnetic resonance imaging (fMRI) of tonotopicity and stimulus intensity coding in human primary auditory cortex. *Laryngoscope*, 114, 512-518.

Zerlin, S. & Naunton, R. F. (1974). Early and late averaged electroencephalic responses at low sensation levels. *Audiology*, 13, 366-378.

Zwislocki, J., Kruger, B., Miller, J. D., Niemoeller, A. F., Shaw, E. A., & Studebaker, G. (1988). Earphones in audiometry. *Journal Acoustical Society of America*, 83, 1688-1689.

Appendix 1: Record CAEPs in infants with success

The general techniques and strategies used in paediatric audiology are highly relevant to the recording of CAEPs in infants. Experienced clinicians will be well aware of how to best manage the test environment, and will have developed many of their own solutions to overcoming the issues that inevitably arise when working with infants.

The clinical study in which CAEPs were recorded simultaneously from HEARLab and Neuroscan TM provided a valuable opportunity to investigate different test techniques, and to consider their relative merits, without the usual restrictions of a routine clinical assessment appointment, in which a diagnostic outcome must be achieved within a limited time frame.

Some general suggestions, based on observations made during the study, are summarised in this Appendix. It is hoped that this information may provide useful guidance, particularly for clinicians with less experience in paediatric audiology, or those who are new to recording CAEPs in young children.

Before the appointment

- Prior to the appointment, provide parents/carers with information (written and verbal) about the procedure. This will reduce the time spent in explanation at the assessment.
- When arranging the appointment, ask about the child's routines. Try to book the test at a time of day when the child is likely to be in a "good" mood, and less likely to be overtired and irritable. Allow plenty of time so that appointment is not rushed, and it is possible to take breaks if needed.
- Check whether the parent/carer intends to bring sibling/s to the appointment. If it is necessary for siblings to attend, make sure there will be suitable activities/supervision away from the test room. The parent should be free to focus their attention on the child having the test.
- Ask the parent/carer to bring food, drinks, or dummies ("pacifiers") for the child, to the appointment. Some favourite toys, that are suitable as quiet distractors, can also be useful in making the child feel more secure in the test environment. DVDs that the child enjoys can also provide familiarity and useful distraction.

- Suggest that the child be dressed for the assessment in layers of clothing that can be easily removed. Electrode contact can be comprised if the child becomes overheated and “sweaty”, and it may be necessary to remove clothing to cool them down. It is better not to have to pull clothes over the child’s head once they are “wired up”.
- Ask parents/carers to ring and postpone the appointment if their child is unwell, particularly if the child has a temperature. A restless and irritable state is not conducive to quality recordings.
- Call the parent/carer to confirm the appointment the day before, and take the opportunity to check whether they have any questions or concerns they would like to discuss.

Test environment

- Make the test environment “child friendly”. For example, decorate the test booth and surrounding areas using items such as mobiles, displays of soft toys (out of the child’s reach), and fabric motifs. Minimise technical “clutter”. Keep wires out of view and laboratory supplies in drawers. Children, particularly if they have undergone medical treatment or hospitalisation in the past, may associate such items with unpleasant procedures.
- Provide a chair that is as large enough for the child to sit comfortably, either on their parent’s lap or beside them, during testing. Some children become irritable if they feel overly restrained or restricted in their movements. Where possible, try to let them settle into a position which they prefer.
- Use washable covers on the chair (eg, bath towelling) and change between assessments to maintain hygiene. This makes “dribbles” and food spills easier to contain and reduces parents inclination to perform immediate “clean ups”, which can disrupt the testing. Have tissues or baby wipes at hand if needed.
- Some younger infants may be comfortable in a rocker/”Fraser chair”, but don’t rock the baby during recordings. “Bounces” can be evident in recordings and rocking can make a child sleepy.
- Fluorescent lighting can cause problems of electrical interference. Incandescent lighting should be used in preference. As well as providing a technical

advantage, incandescent lamps can create a pleasant, relaxed ambience for children and parents. Novelty lamps (eg, artificial fish tanks, “lava” lamps), as well as providing illumination, can provide the child with visual distraction.

- It is important that the tester is able to monitor the test environment and the baby’s state throughout the test. A strategically placed video camera can be extremely helpful if the arrangement of the test booth makes it difficult to maintain a clear view.
- Where the tester is in a separate observation room, an audio monitor is essential in monitoring the ambient noise level in the test environment, and is also useful in communicating with the distractor.

Preparation for testing

- Children generally have a short attention span, and their mood and state can change quickly. Have all test equipment switched on, checked and calibrated before the child arrives. Have the recording system software open, and ready on the impedance check screen to avoid unnecessary delays.
- If testing is with hearing aids on, change the batteries and check the devices on arrival. Having another staff member do this, while you are interviewing the parent, will minimise delay in commencing testing.
- Ensure that the parent understands what is involved in the test. The child is more likely to be relaxed and cooperative during the assessment if their parent is confident and relaxed about what is happening.
- Make the parent comfortable. Providing a hot drink or glass of water can help put parents at ease.
- Ensure the parent feels in control of the situation. Seek their advice about the child’s preferences, and the best strategies for preparing them for testing. Respect their opinion and follow their suggestions.
- Try to build some rapport with the child. A little physical interaction with the child while you are interviewing the parent (eg, patting or stroking their arm or head) can be useful in gauging how they will react to the preparation for electrode sites, and may possibly help the child accept it more readily.
- Make sure the child’s physical needs (eg, nappy changing) are attended to before starting preparation for the electrode placement.

- Have any items that might be needed (eg, bottles, food, toys) close at hand, in order to minimise noise and disruption during testing. Check with parents to see if bottles need warming or food needs preparation before you start.
- Attempt otoscopy and tympanometry first if indicated, but don't persevere if it causes the child to become too active or distressed.
- Ask parents to switch off mobile 'phones or pagers as these may cause distraction to the child if they ring during testing, and may also be a potential source of electrical interference.

Preparation of electrode sites

- Attaching the electrodes is potentially the most challenging part of the test procedure. Approach the preparation confidently but not too forcefully. Smile, and talk to the child reassuringly.
- Start the preparation in a position where the child is comfortable and not overly restrained. For example, try starting while an older infant is playing on the floor, or at a child's table and chair.
- Try not to physically "stand over" the child while doing skin preparation. Working from behind the child may be a good option.
- Electrode sites are generally prepared by abrading with a cotton applicator 'bud' and a medical gel intended for the purpose. Rub firmly and vigorously enough to cause a slight redness on the skin surface, but not so hard that the child becomes obviously distressed by the sensation.
- Work as quickly as possible and minimise the number of physical contacts with the child. Don't fuss or "overdo" it, but be mindful that it is better to prepare skin thoroughly than to have to repeat the whole process.
- If the child needs reassurance about the preparation, modelling the procedure (eg, by rubbing the forehead of the parent or a doll with a cotton bud, and sticking on an electrode) can be helpful. Try letting an older child have a "turn" at putting an electrode on a toy or on their parent.
- Some children will be reassured by watching the preparation in a mirror, but this may make others more apprehensive.

- Television can be a good distraction during electrode placement. Use a range of children's DVDs with lots of colour and movement. If the child becomes interested, the DVD can be left playing (with the sound muted) when testing starts.
- Wherever possible have a trained distractor (in addition to the parent), to interact with the child during electrode placement, as well as during testing. Toys that involve some fine motor manipulation (eg, block stacking, button pressing) can help keep hands away from the electrode sites, and at this stage toys that make a noise are suitable.
- Cleaning skin with an isopropyl alcohol prep swab after abrading is sometimes recommended, and may improve contact, but it can make the electrode stick very firmly and make it difficult to remove. Therefore, alcohol preparation is not recommended for the delicate skin of infants.

Optimising and maintaining electrode contact

- Use a liberal amount of electrode paste under the vertex electrode even if a disposable electrode, that already contains conductive gel, is used.
- If using disposable electrodes, a spot of double sided tape (the type used for retaining hearing aids) on the underside of the plastic tab of the electrode stud, can give a firmer hold, particularly for mastoid or forehead sites (i.e., where the skin is free of hair).
- A headband is very helpful in keeping the electrodes in place, particularly at the vertex, but some children are less accepting of wearing a headband than others.
- To make wearing a headband more appealing to the child, choose colourful, soft and stretchy materials. Give older children a choice of colours or designs (eg, have a selection of different motifs sewn on a selection of headbands). Having a choice of "girls" or "boys" styles can be important to the child, and sometimes also to the parent. Use fabrics that are easy to wash and dry after use.
- Dividing the top of the headband before use (by cutting a slit a few inches long across its centre) allows a section of fabric to be stretched forward to hold the forehead electrode in place.

- Passing the leads under the headband can help reduce pulling and strain on the electrode site during the test.
- Elastic bandage, particularly of material that allows the ends to adhere without pins or tape (eg, “peg” bandage), can be a reasonable alternative, but tends to be more “fiddly” to put on than a headband. If an inexpensive type is chosen it can also be disposed of after use which can be an advantage.
- Once the electrodes are attached, try to drape the leads behind the child, avoiding contact with the child's face or neck. If the child can feel them, they will be more inclined to pull at them. Try to keep the leads from becoming tangled in bibs or collars.
- Loosely taping the leads to the back of the child’s clothing (using micropore tape) may be helpful, but ensure they are not taped so tightly that the electrode leads are pulled off the head if the child suddenly leans forward.
- Make sure that if the parent is holding the child, that the electrode leads are not cramped or pulled under the parent's arm. Try directing the leads up and over the parent’s shoulder
- Avoid the child leaning back onto the electrode leads, or making sudden large movements, such as lunging forward. Strategic use of distraction toys can help in this respect.
- If the child starts to touch the leads or electrodes don't over-react (eg, grab suddenly at the child’s hands). In preference, try to distract the child by offering them an alternative item to play with.
- Once the electrodes are in place, avoid touching them unless really necessary (ie, they are obviously slipping/becoming unstuck). Drawing child's attention to them will often result in renewed efforts to remove them.
- Don't let the child overheat. This can result in electrodes lifting and the reject rate increasing. If the child gets very restless it can be better to suspend testing than let it proceed until they are “hot and bothered”.

Distraction techniques

- The distractor is best seated in a comfortable position at, or below, the child's eye level. Care must be taken to maintain an appropriate position in relation to the speaker when the stimulus is presented free-field.
- Have a wide selection of age appropriate toys that are not too noisy. Keep toys and other distraction aids in easy reach of adults, to minimise noise and disruption as the test proceeds. Choose toys that can be cleaned according to infection control procedures. Toys that can't be cleaned (eg, soft toys) should be kept out of the child's reach.
- For very young infants, the main aim of distraction is to keep them alert and awake as they obviously don't have the motor skills to pull at the leads or electrodes. Mobiles, hand and finger puppets are all useful items. Visual novelties (eg, toys with lights and motion) can be excellent. Some mechanised toys are too noisy to use while the testing is in progress, but can be good to use during breaks, in order to regain a child's interest and increase alertness.
- Toys that are used for VROA distraction are generally suitable for children in the 7- 24 month age group. Examples include; stacking plastic rings or cups, farm animals, large counting and threading beads, puzzles, colourful teething rings and so on. Items that keep hands occupied are ideal for children old enough to manage them.
- For younger children or those who are not developmentally ready to "play" with items, toys with texture (eg, spiky plastic balls, plastic animals etc) can be interesting for the child to touch or mouth.
- Action toys (eg, water-wheels, small spinning tops, "pecking" birds, "wobbly" animals, clear plastic balls that contain moving toys) can be useful as long as they are not too noisy. If children are allowed to hold the items they should not contain small parts that may be a choking hazard, and they must be easy to clean. Watch that water-filled toys don't leak.
- For older infants (around 24 months and over) try colouring-in books with big colourful crayons, play-dough (if past eating it, or well supervised), paper with stickers or self-inking stamps.

- Books can be good for children of all ages. Heavy cardboard books of various shapes, or with flaps/pop-up features, can provide “hands on” activity. Books made of plastic (ie, intended for bath time) can be ideal for very young children as they are easy to clean if mouthed.
- Almost every child enjoys watching soap bubbles being blown. The small bottles used for parties are inexpensive and easy to use. Avoid “sticky” bubbles that are designed not to burst. They tend to leave messy residue in the test environment.
- Eating and drinking are excellent distractors but may increase the electrical noise levels within the response. Good choices include baby bottles or infant sipping cups, and soft foods such as banana, custard, fruit gel, or sultanas. Avoid hard foods (eg, crunchy crackers) or large pieces of food of that require a lot of chewing, as the resulting noise and jaw movements will almost certainly affect recordings.
- Breastfeeding is good for calming infants, but often can induce sleep. If the baby must feed during the assessment, watch very carefully and rouse the child gently if they begin to doze off, or it appears that their eyes are becoming “unfocussed”. Be prepared to pause testing if the child’s state becomes inappropriate.
- Avoid distraction activities that are too stimulating or that encourage increased vocalisation, for example, physically vigorous play, or games/gestures that encourage the child to answer questions or name objects. Parents may sometimes need some guidance about activities that are inappropriate in this respect.
- If the child is content and quiet the distractor should withdraw and leave them to their own devices. If the child is unsettled sometimes it can help for the distractor to get right out of their view and let the parent try to calm them.
- Don’t persist with testing if the child is distressed. It is probably more time efficient in the long run to make another appointment.

Appendix 2: Cochlear implants and HEARlab

Objective techniques for evaluation of all device fittings are critical to realising the full benefit of early detection of hearing loss and intervention. Electrophysiological techniques have been routinely used to supplement behavioural measures in cochlear implant evaluation (Shallop, 1993) with CAEPs being recorded from cochlear implant (CI) wearers using both electrically generated stimuli (direct to the implant itself), and various acoustic stimuli.

As well as providing objective validation of device fitting, particularly for younger children, CAEPs have also been used to a) study maturational changes of the central auditory system in CI patients (Gilley, Sharma, Dorman, & Finley, 2006; Ponton, Don, Eggermont, Waring, & Masuda, 1996; Eggermont, Ponton, Don, Waring & Kwong, 1997; Sharma, Dorman, Spahr, & Todd, 2002; Sharma, Dorman, & Spahr, 2002) and b) differentiate between subjects with “good” versus “poor” speech reception performance (Friesen & Tremblay, 2006; Maurer, Collet, Pelster, Truy, & Gallego, 2002).

Characteristics of the CI artifact

When recording CAEPs from participants with CIs an electrical artifact may be evident which can significantly interfere with the identification of a true neural response (Gilley et al., 2006; Martin, Tremblay, & Stapells, 2007). The longer the stimulus duration, the more likely this is to occur (Martin, 2007).

The CI artifact, which generally presents in the averaged CAEP response as a square-wave with a duration at least equal to that of the acoustic stimulus, can be five to ten times larger in magnitude than the averaged evoked response. It is mainly attributed to the radio frequency transmission of signals from the implant transmitter to the receiver, although it is believed that other parts of the implant may also contribute (Martin, 2007).

The characteristics of the artifact appear to depend on internal device factors (e.g., the number of active electrodes in the array, the position of these electrodes in the cochlea

and the location of the return electrode) and /or the individual processor settings (e.g., the number of channels, stimulation rate, pulse rate , pulse duration and maxima).

The morphology of the P1-N1-P2 complex is affected by maturation (Martin et al, 2007) with responses typically falling within the 50 to 250 ms range in adults (Tremblay & Burkard, 2007). For newborn infants however, the latency of the P1 can be as long as 300-400 ms (Sharma, Dorman, & Spahr, 2002; Gilley et al, 2006). Where the region of interest for detection of a response is beyond the range of the stimulus onset and offset, it could be assumed that the artifact would be less of an issue but overshoots and subsequent “ringing” of the recording amplifiers may increase the duration of the artifact beyond 400 ms (Gilley et al., 2006). In addition, a low magnitude noise floor, attributed to the ongoing background stimulation by the speech processor, has been observed throughout the entire recording period (Gilley et al., 2006).

HEARLab evaluation trials : Two case studies

Given the potential application of CAEP recording for CI evaluation, it was considered appropriate to include some observations from CI wearers as part of the *Clinical Evidence for the Reliability and Validity of HEARLab Module 1* study. Two adults with CIs were assessed using the ACA Module of HEARLab. Although they were not recruited according to any specific criteria, they were well matched in terms of age, onset of hearing loss, length of experience with a CI, and device and processor type.

Case 1

Case 1 was a 62 year old female who had received a CI to her right ear in 1998. The onset of her hearing loss was approximately 22 years prior to implant. At the time of HEARLab testing, she was reportedly working with her audiologist to try and obtain a better ‘map’ with her current speech processor. She felt that she had experienced better performance in the past with a different processor and ‘map’, however she was generally positive about the use of her CI, even with the current parameters.

Processor type: Freedom

Mode- MP1 (monopolar)

Map 1	Strategy	Rate	Maxima	Pulse width	# channels	Implant
1	ACE (Beam + ADRO + ASC)	500	8	37	1-22	C124M
2	ACE (Beam + ASC)	500	8	37	1-22	C124M
3	ACE (ADRO)	500	8	37	1-22	C124M
4	ACE (ADRO+35dB11DR)	500	8	37	1-22	C124M

CAEPs were recorded using the speech stimuli /t/ and /g/ presented at 65 dB SPL in the free-field. The CI processor was left on the default sensitivity setting.

Map1:

Figure 1 shows a rhythmic sinusoid-type pattern in HEARLab's ongoing eeg trace which was repeatable on a second run. With the test partially completed, the averaged waveform did not show any evidence of CI artifact. Figure 2 shows the final averaged CAEP for the stimulus /g/ with a detection P value of 0.001, a convincing CAEP response of normal latency and no apparent CI artifact.

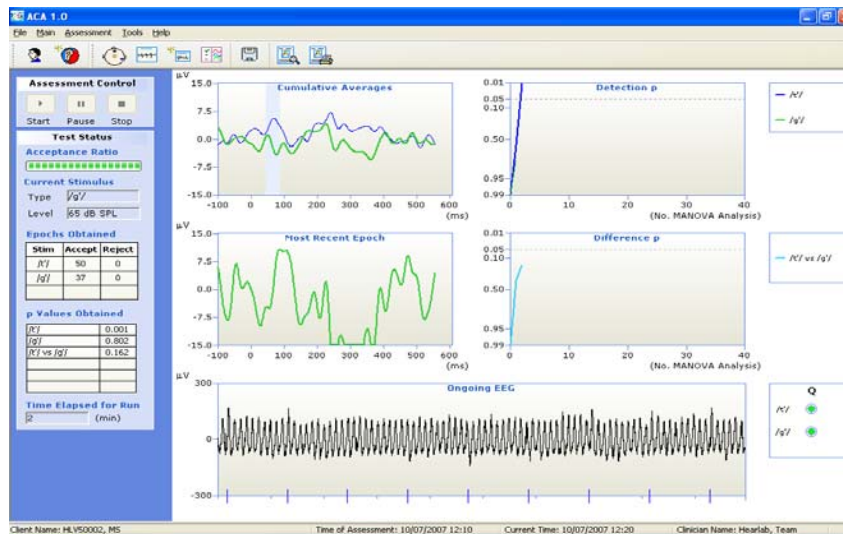


Figure 1: Case 1, Map 1 recording in progress

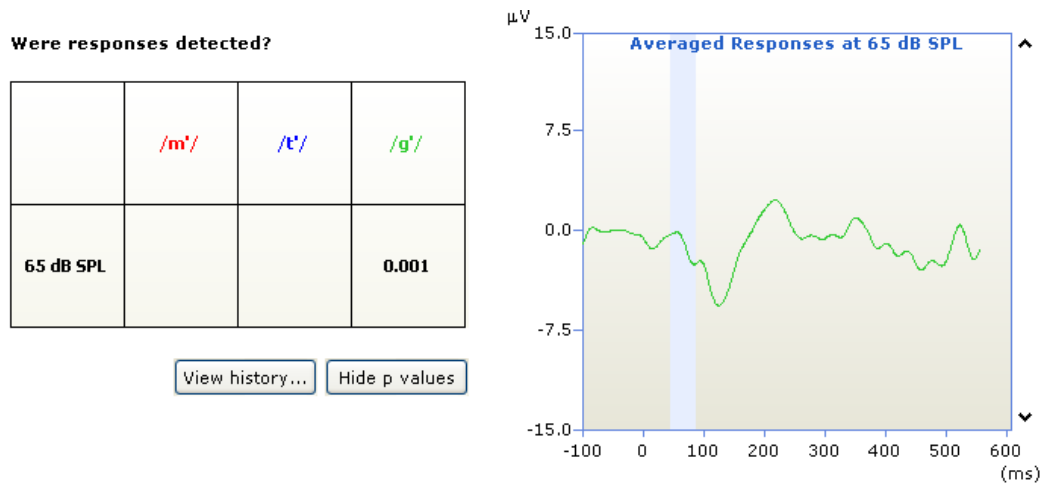


Figure 2: Case 1, Map 1 second (repeat) recording with 146 accepted epochs (nil rejected)

Map 2, 3 & 4:

A single run for the stimulus /g/ was recorded for each of the remaining maps. A similar sinusoid-type pattern was evident in the ongoing eeg of each recording, however its appearance was slightly different for each Map setting as shown in Figures 3, 4 & 5, and different to Map 1 as shown in Figure 1. The morphology of the averaged responses however appears to be quite acceptable (see Figures 4 & 5).

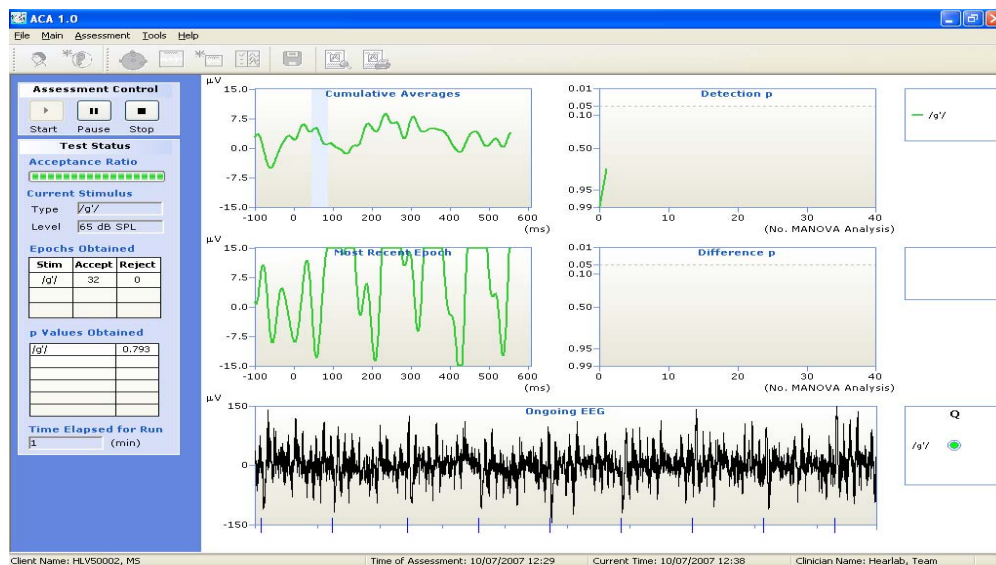


Figure 3: Case 1, Map 2, recording in progress

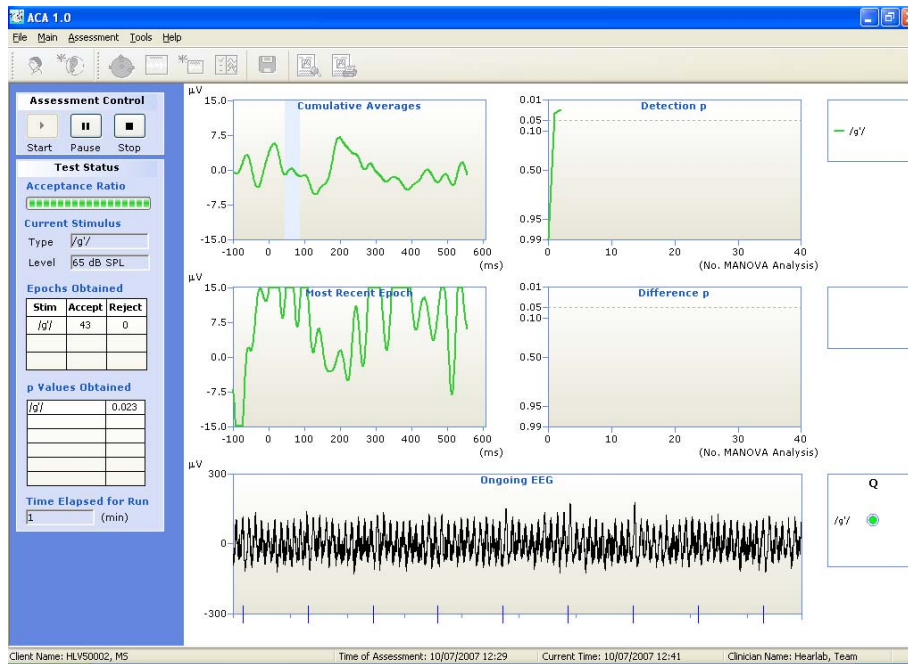


Figure 4: Case 1, Map 3, recording in progress

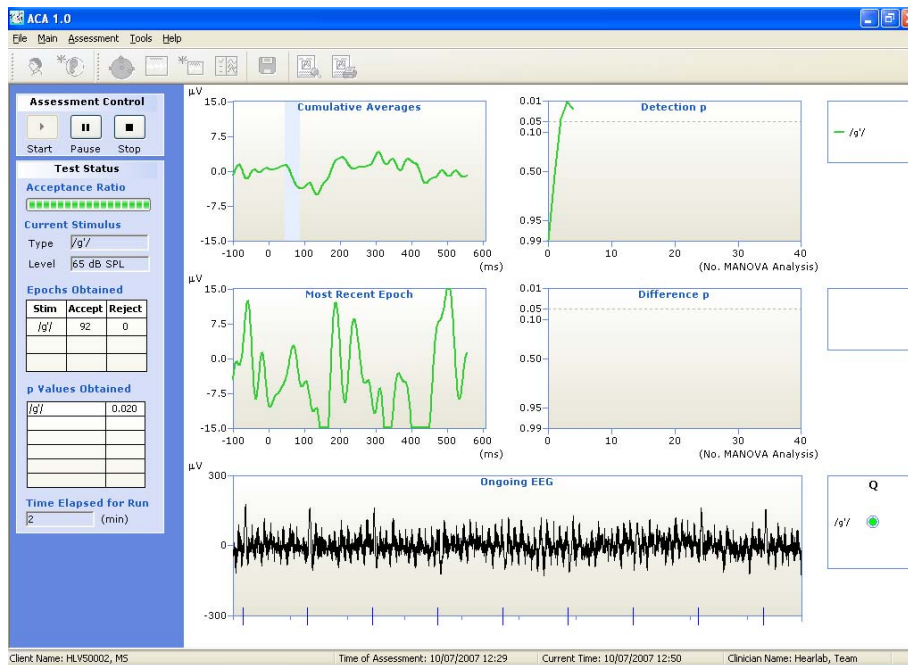


Figure 5: Case 1, Map 4, recording near completion with 92 accepted epochs.

To check that the “interference” was not originating from HEARlab’s amplifiers, an additional run was recorded (not shown) using the /t/ and /g/ stimuli with recording electrodes on a resistor pad rather than the participant’s head. Under this condition no “interference” was observed.

Conclusions: The CI was clearly interfering with the normal recording of the ongoing eeg but the effect differed as each Map was activated. These maps reflect different processing strategies but were otherwise identical. The final averaged CAEP traces did not appear to be significantly affected by stimulus artifact.

Case 2

Case 2 was a 57 year old female who received her CI in the right ear in 1997. The onset of the hearing loss was approximately 23 years prior to implantation. Satisfaction with the performance of the CI was somewhat limited (according to the wearer and their audiologist).

Processor type: Freedom

Mode- MP1 & 2 (Monopolar)

Map	Strategy	Rate	Maxima	Pulse width	#channels	Implant
1	ACE	900	6	25	7-22	C124M
3	ACE	250	8	25	7-22	C124M

(NB; Map 2 and 4 had the same paramaters as Map 1 and 3 respectively, but with the addition of a noise reduction feature).

CAEPs were recorded using HEARLab’s interleaved speech stimuli /m/, /t/ and /g/ presented at 65 dB SPL in the free-field. The CI processor was left on the default sensitivity setting (“5”).

Map 1:

HEARLab’s ongoing eeg trace showed a regular periodic “pulse” throughout the recording that was repeatable on a second run. Figure 6 shows the final averaged responses to one of these runs. There appears to be a large amplitude negative-

deflection with a time period of 10 - 80 ms approximately which is consistent with descriptions of a CI artifact. This artifact overlaps in time with the duration of the speech stimuli and with the expected P1-N1 response. The CAEP detection P values, which use an analysis period commencing at 76 ms in adults, also overlap with this time period and hence a low (i.e, significant) P value is returned. Given the presence of the apparent CI stimulus artifact, this statistical outcome is misleading.

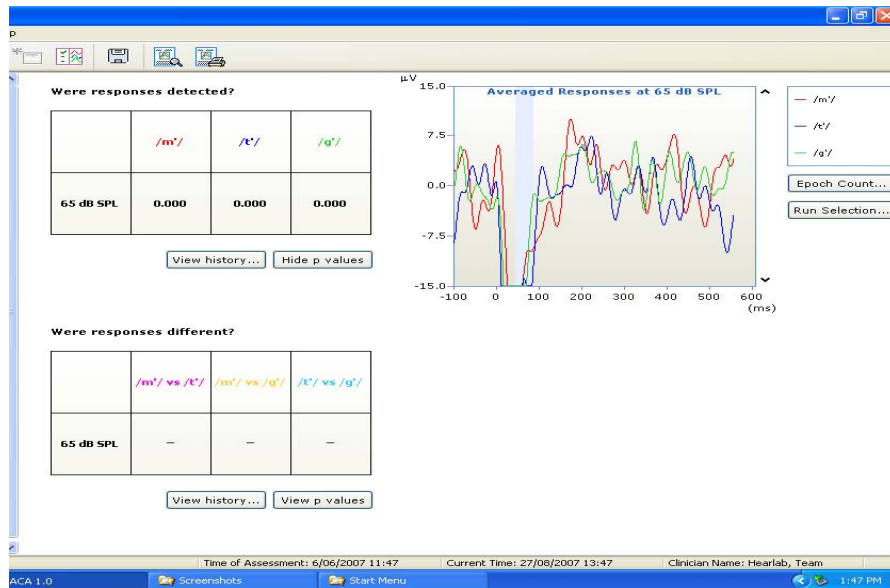


Figure 6: Case 2, Map 1, with 75 – 85 accepted epochs for each stimulus generated response.

Map 3:

With the CI set to map 3, HEARLab's ongoing eeg again showed a regular periodic "pulse" throughout the recording. Figure 7 shows the final averaged recording for this Map with no evidence of a negative-deflection and no apparent CAEP. The detection P values for all three speech stimuli were high suggesting no CAEPs were evident.

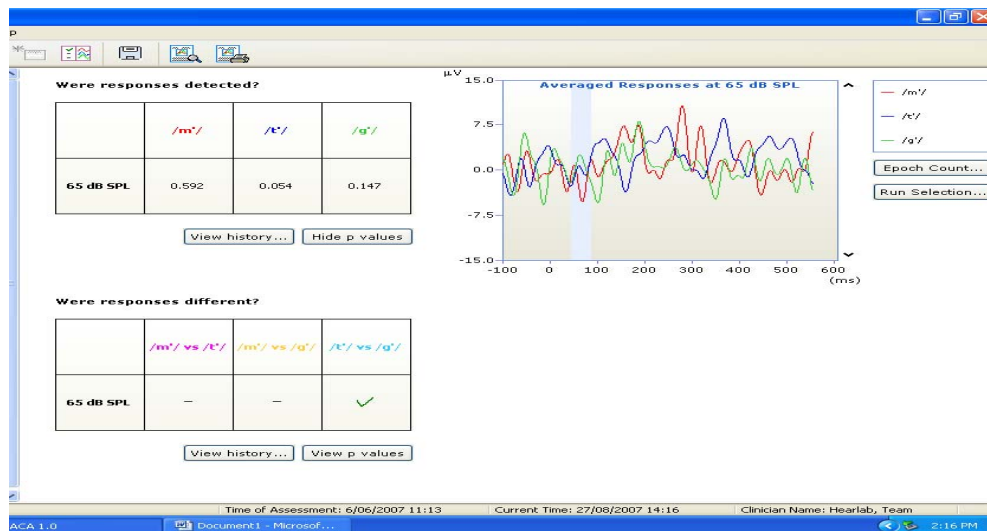


Figure 7: case 2, Map 3, recording complete with 70- 90 accepted epochs for each stimulus-generated response.

To check that the “interference” was not originating from HEARlab, two tests were performed. Firstly, the speech processor was switched to the Map 1 program but the loudspeaker was disconnected to provide a non-stimulus trial. No “pulse” was evident in the ongoing eeg, the detection P value was not significant and no CAEP could be observed. Secondly, the speech stimuli were presented and recordings were made while the participant’s speech processor was switched off. Again, there was no observed “pulse” in the ongoing eeg, the detection P value was not significant and CAEPs could not be observed. It can be concluded that the artifact observed when the CI was switched to Map 1 had its origins in the CI and not in HEARLab.

Conclusions: In this case, a regular periodic “pulse” was evident in HEARLab’s ongoing eeg when the CI was switched to Map 1 and Map 3. An artifact was however evident in the final averaged response using Map 1 only. The only differences between the two Maps were a) the stimulation rate which was faster in Map 1 than Map 3, and b) a difference in the maxima setting.

Recommendations:

It is acknowledged that the two case studies presented here are a very limited sample. It appears however that artifacts are clearly evident in the averaged response with some CI processor settings and not with others. In these examples, the artifact was

observed with the highest stimulation rate and lowest maxima. A systematic trial that investigates the contribution of various processing parameters to the generation of the artifact is therefore recommended. If HEARLab is used to record CAEPs in CI cases, it is important to reiterate that reliance on the detection and difference P values without due consideration of the averaged wave morphology may be misleading.

The following recommendations are made when using the NAL-ACA module with CI wearers:

- The ongoing eeg should be carefully observed for unusual “pulses” however their existence may not necessarily mean that averaged responses will be invalid.
- The detection and difference P values should be examined in the light of the final average responses. If any unusual peaks or troughs are evident, particularly within the first 100 ms after stimulus onset, the results may be invalid.
- Stimulus artifact may have effects beyond the first 100 ms post stimulus onset. Traces with unusual morphology should therefore be interpreted with caution.
- Where contamination by CI artifact seems likely, clinical decisions should not be based on the results of the CAEP tests.

References

- Eggermont, JJ., Ponton, CW., Don, M., Waring, MD., and Kwong, B. (1997). Maturational Delays in Cortical Evoked Potentials in Cochlear Implant Users. *Acta oto-laryngologica*, 117: 161-163.
- Friesen, LM., & Tremblay, KL. (2006) Acoustic Change Complexes Recorded in Adult Cochlear Implant Listeners, *Ear & Hearing*, Vol. 27, No. 6, 678-684.
- Gilley, PM., Sharma, A., Dorman, M., Finley, CC., Panch, AS., Martin, K. (2006). Minimization of cochlear implant stimulus artifact in cortical evoked potentials, *Clinical Neurophysiology*, 117, 1772-1782
- Martin, BA. (2007). Can the Acoustic Change Complex Be Recorded in an Individual with a Cochlear Implant? Separating Neural Responses from Cochlear Implant Artifact, *Journal of the American Academy of Audiology*, 18: 126-140.
- Martin, BA., Tremblay, KL., Stapells, DR. (2007). Principles and Applications of Cortical Evoked Potentials. In Burkard RF., Don M., Eggermont JJ. (Eds) *Auditory Evoked Potentials Basic Principles and Clinical Application*. Lippincott Williams & Wilkins.
- Maurer, J., Collet, MD., Pelster, H., Truy, MD., Gallégo, S. (2002) Auditory Late Cortical Response and Speech Recognition in Digisonic Cochlear Implant Users, *The Laryngoscope*, 112: 2220-2224.
- Ponton, CW., Don, M., Eggermont, JJ., Waring, MD., and Masuda, A. (1996) Maturation of Human Cortical Auditory Function: Differences Between Normal-Hearing Children and Children with Cochlear Implants, *Ear & Hearing*, Vol.17, No. 5, 430-437.

Shallop, JK. (1993) Objective Electrophysiological Measures from Cochlear Implant Patients, *Ear & Hearing*, Vol. 14, No. 1, 58-63.

Sharma, A., Dorman M., Spahr, A., Todd, NW. (2002a) Early Cochlear Implantation Allows Normal Development of Central Auditory Pathways *The Annals of otology, rhinology and laryngology* 111, 38-41,

Sharma, A., Dorman, MF., Spahr, AJ. (2002b) A Sensitive Period for the Development of the Central Auditory System in Children with Cochlear Implants: Implications for Age of Implantation, *Ear & Hearing*, Vol. 23, No. 6, 532-539.

Tremblay, KL. & Burkard, RF. (2007). Confounding Effects of Hearing Loss on AEP's. In Burkard RF., Don M., Eggermont JJ. (Eds) *Auditory Evoked Potentials Basic Principles and Clinical Application*. Lippincott Williams & Wilkins.