Fully objective hearing threshold estimation in cochlear implant users using phase-locking value growth functions

Darren Mao[a,b], Hamish Innes-Brown[b,c], Matthew A. Petoe[b,c], Yan T. Wong[a,d], Colette M. McKay[b,c]

[a] Department of Biomedical Engineering, University of Melbourne, Parkville VIC 3010, Australia

[b] The Bionics Institute, 384-388 Albert St, East Melbourne VIC 3002, Australia

[c] Department of Medical Bionics, University of Melbourne, Parkville, VIC 3010, Australia

[d] Department of Physiology, Department of Electrical and Computer Systems Engineering, and the Biomedicine Discovery Institute, Monash University, Clayton VIC 3168, Australia

Corresponding author: Darren Mao, DMao@bionicsinstitute.org

Abstract

Cochlear implant users require fitting of electrical threshold and comfort levels for optimal access to sound. In this study, we used single-channel cortical auditory evoked responses (CAEPs) obtained from 20 participants using a Nucleus device. A fully objective method to estimate threshold levels was developed, using growth function fitting and the peak phase-locking value feature. Results demonstrated that growth function fitting is a viable method for estimating threshold levels in cochlear implant users, with a strong correlation ($r = 0.979$, $p < 0.001$) with behavioral thresholds. Additionally, we compared the threshold estimates using CAEPs acquired from a standard montage (Cz to mastoid) against using a montage of recording channels near the cochlear implant, simulating recording from the device itself. The correlation between estimated and behavioural thresholds remained strong ($r = 0.966$, $p < 0.001$), however the recording time needed to be increased to produce a similar estimate accuracy. Finally, a method for estimating comfort levels was investigated, and showed that the comfort level estimates were mildly correlated with behavioral comfort levels ($r = 0.50$, $p = 0.024$).

Keywords

Cochlear implant, electrical dynamic range, EEG, objective threshold estimation, phase-locking value, cortical response.

1 Introduction

Cochlear implant stimulation needs to be adjusted for each individual in order to provide the appropriate range of sound perception for optimal speech understanding. An important part of the process involves determining the electrical dynamic range. The lower bound of the dynamic range corresponds to a stimulus that produces a just-perceivable sound and is termed the threshold (T) level; the upper bound corresponds to a stimulus that produces a comfortably loud sound and is termed the comfort (C) level. Clinically, the dynamic range is usually determined by behavioral feedback from the implantee. This process is time consuming and unfeasible for infants (Hughes, 2013) and other individuals who are unable to give their own feedback or follow instructions. Objective determination of the dynamic range is advantageous when behavioral feedback is not available. Thus, much research has been undertaken to evaluate whether neural responses to stimulation can be used to determine objective estimates of threshold and comfort levels. However, current objective measures may be inaccurate, time-consuming or too labour or resource intensive to have widespread clinical utilization (Vaerenberg et al., 2014). A less-explored, but promising, objective measure is the cortical auditory response (CAEP, Visram et al. (2015)), along with the phase locking value feature (Mao et al., 2018). The present study used phase-locking values extracted from single channel cortical responses to estimate the threshold level for cochlear implant users.

A widely investigated neural response is the electrically-evoked compound action potential (ECAP). ECAP thresholds only have a moderate correlation with behavioural thresholds (a maximum *r* value around 0.5 to 0.6, which decreases as rate of stimulation of the behavioral measure increases) (Abbas and Brown, 2015; McKay and Smale, 2017; McKay et al., 2013; Miller et al., 2008). Despite its limited accuracy, ECAPs are used as a rough guideline for clinical implant fitting due to easy acquisition with neural telemetry hardware incorporated in modern implants (Shapiro, 2014). Additionally, ECAPs use single stimulus pulses, which is not representative of the stimuli used for clinical threshold finding (high-rate pulse trains). The electrically-evoked auditory brainstem response (eABR) suffers from the same problem. Only single or short pulses can be delivered as to not obscure the response with stimulation artefact. Thus, low rates of stimuli generate relatively good correlations with behavioral threshold (for example, $r = 0.89$ at 10 pps (Hodges et al., 1994) and $r = 0.83$ at 35-80 pps (Brown et al., 2000)). At higher stimulation rates that are more representative of what is usually used in clinic, eABR thresholds are only marginally more accurate than ECAP thresholds (r = 0.74 at 400 pps, r = 0.69 at 1000 pps (Brown et al., 1999)). In contrast, using the electrically-evoked auditory steady state response (eASSR) with seven channel and 96 kHz sampling rate, Hofmann, Wouters (2010) found thresholds that were highly correlated with behavioral thresholds (r = 0.96), although one of 26 subjects' thresholds could not be estimated. The hardware requirement for this method is currently a barrier in its translation to clinical usefulness; in addition, paediatric testing with eASSRs require the patient to be still, which also limits its clinical usefulness.

Cortical responses as an objective measure for cochlear implant threshold estimation is a developing area of research. Visram et al. (2015) showed high correlations (r = 0.93) between behavioral thresholds and cortical thresholds obtained using high-rate pulse train stimuli and using the global field power feature. The global field power is found by calculating the standard deviation of EEG amplitude across every recording channel. The requirement for multiple recording locations is not ideal for easy clinical use. Another challenge for automatic cortical evoked response estimation is the way that morphology changes with age, particularly in the infant population (Kushnerenko et al., 2002; Silva et al., 2013; Wunderlich et al., 2006). In our previous work using cortical responses in subjects with normal acoustic hearing (Mao et al., 2018), we presented an algorithm to address the challenges of both ease of clinical use and morphology heterogeneity: we used single channel acquisition, which allows easier translation into clinical situations; the algorithm also made no assumptions about the shape of the response morphology. The present work examined whether a similar method can effectively estimate electric hearing thresholds in cochlear implant users, and thus its suitability for automatic threshold determination in cochlear implant clinics.

Traditionally, cortical responses are recorded from the Cz or nearby positions (Abbas and Brown, 2015; Cone and Whitaker, 2013). Studies have shown altered morphology and reduction in cortical response amplitude when recording location is varied, which is likely due to sub-optimal alignment of the EEG electrodes with the electric dipoles generated by responding neurons (Kelly et al., 2005; Näätänen and Picton, 1987). However, it is unknown whether the phase-locking value (PLV) of the cortical response, which quantifies the similarity in phase across response trials of the same stimulus (Delorme and Makeig, 2004; Tallon-Baudry et al., 1996), is affected similarly, as its use in the hearing field is limited. In this study, we investigated whether threshold estimation using the phase-locking value feature of the cortical auditory potential is robust to changes in electrode location. Because the method in this paper may be applicable to future technology in which implanted electrodes of a cochlear implant are used to measure the cortical evoked response, we investigated how the recording location of the EEG response affects the accuracy of threshold determination.

In summary, we hypothesized that the peak phase-locking value feature of the cortical auditory potential would allow accurate estimation of electrical hearing thresholds in cochlear implant users. In addition, we compared the deterioration in amplitude and phase-locking value when recording location was changed to a more clinically convenient location, and hypothesized that the correlations between the objective estimates using phase-locking value and behavioral measures would be still clinically significant when the data quality was deteriorated by recording from a suboptimal recording position. Finally, we investigated whether comfort levels could be estimated using the PLV feature growth functions.

2 Methods

2.1 Subjects and stimuli

20 adult users of cochlear implants participated in this study. Their ages ranged from 39 to 86 (mean = 65.2, std. = 10.7) years. All implants were Nucleus CI24RE implants and had been in use for at least three years. Table 1 shows the demographic information for each subject. This study was conducted under the approval of the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital, and each participant gave written informed consent.

The test stimulus for both behavioral and EEG measures activated the implant electrode assigned to the 1kHz frequency, which was electrode 16 for all subjects except one (for whom it was electrode 14). Test stimuli were 50-ms duration biphasic pulse trains delivered to this electrode in monopolar 1+2 stimulation mode. The current level was varied trial-to-trial to create stimuli of different intensities. Pulse widths of 26 µs, interphase gap of 8.4 µs, and rate of 900 pulses-per-second were used for 15 out of 20 subjects. The remaining five subjects needed a greater pulse width or pulse rate to reach comfortable loudness without exceeding the Nucleus devices' current level upper limit of 255 units(shown in Table 2).

Table 1. **Subject demographic information and implant side.**

Subject number	**Age**	**Gender**	**Duration of CI use**	**Test ear**	**Cause of deafness**
1	70	F	Unknown	R	Measles when young
2	67	M	5	R	Otosclerosis
3	68	F	3	L	Otosclerosis
4	39	F	4	R	Unknown
5	69	M	6	L	Hereditary
6	71	F	3	L	Hereditary
7	74	M	1.5	L	Injury
8	48	M	8	L	Congenital
9	58	M	10	L	Injury
10	59	F	6	L	Unknown sudden onset
11	52	M	6	R	Meningitis
12	61	M	8	R	Congenital

13	62	M	4	L	Hereditary
14	64	M	1	R	Unknown progressive
15	86	F	7	L	Ototoxicity
16	76	M	Unknown	R	Progressive hereditary
17	78	M	3.5	R	Unknown
18	65	F	3	R	Unknown
19	71	M	4	R	Unknown sudden onset
20	66	F	8	R	Unknown

Table 2. The stimulation parameters pulse width and pulses-per-second used for each subject.

Subject number	**Test pulse width (μs)**	**Test pulses-per-second**
2	50	900
3	62	900
5	62	900
6	50	900
12	26	1800
Other 15 subjects	26	900

A 50-ms stimulus duration was used for cortical measurements, as it is the longest duration that does not obscure the cortical response with the cochlear implant stimulation artefact (Visram et al., 2015). The subjects' clinical maps were created using 500-ms duration pulse train stimuli, which may be perceived as louder than a 50-ms pulse train of the same current level (Donaldson et al., 1997). Therefore, we re-determined the comfort and threshold levels of all subjects using the same 50-ms pulse trains using the following behavioral procedures.

2.2 Behavioral data collection

Behavioral comfort levels were determined using a categorical loudness scale, which had levels of “unheard”, “very soft”, “soft”, “soft but comfortable”, “comfortable”, “loud but comfortable”, “very loud” and “too loud”. The stimuli were presented sequentially, beginning at a current level near each subject's clinical map threshold level, and increased by ten current levels until "comfortable" category was reached, then by five current levels until "loud but comfortable" was reached and finally by two current levels until "very loud" was reached. The current level was then reduced in two current level steps until "loud but comfortable" was reached again, and this final current level was deemed the behavioral comfort level. The behavioral comfort level is defined in this paper as the upper bound of the dynamic range (DR), also referred to as the behavioral 100% DR level.

Behavioral threshold levels were determined using an adaptive, 2-down 1-up, three-interval, three alternative, forced-choice task using the ImpResS research interface (a custom-built system for psychophysics using direct electrical stimulation). Each of the three intervals were 1000 ms in duration, and only one randomly chosen interval contained the 50 ms pulse train. Subjects were asked to select the interval that contained sound and to guess if unsure. Each interval lit up with a button on the ImpResS research interface. Stimuli were initiated at the "soft but comfortable" level. The step size was four current levels for two turns and then two current levels for eight further turns. The task was repeated and the average of the two staircases' final six turning points was deemed to be the behavioral threshold level, or behavioral 0% DR level.

2.3 EEG data collection

Electrical stimuli were presented at six levels relative to the subjects' behavioral dynamic ranges: -50, 10, 20, 40, 60 and 100% DR. The EEG data from the -50% DR sub-threshold stimuli were used as baseline levels in the analysis. Stimuli were presented with an inter-stimulus interval randomly jittered between 1.35 and 1.65 seconds, and each of the six stimulation levels were presented in a random order for at least 300 epochs per level depending on testing time limitations. EEG data were acquired with a BioSemi ActiveTwo EEG system from 64 scalp channels arranged in the international 10-20 configuration, with the Common Mode Sense (CMS) and Driven Right Leg (DRL) placed on the left and right sides of the midpoint between Pz and POz respectively. Left and right mastoid channels were also included.

Data were sampled at 2048 Hz, and a fifth order sinc response anti-aliasing filter with -3dB cut-off at 400 Hz was applied by the BioSemi hardware. Cochlear implant stimulus artefacts, which included the power-up pulses in the 100 ms preceding stimulus onset and the 50 ms of stimulus artefact, were removed before any post-processing by substituting data from a prestimulus region randomly selected from within -600 to -300 ms in each epoch. Substitution was used to ensure the time-frequency spectrogram used to calculate the PLV did not have discontinuities.

For each subject, one to five measurement channels were blocked by the implant's external coil position (average 2.4 channels) and were disconnected. In addition, one to three channels per subject (average 0.6) were too noisy to use and were rejected via visual inspection (visible, very large line noise interference). Montages were created by referencing each unrejected channel to the mastoid channel contralateral to the side of the cochlear implant; each channel is referred to by their location and reference, for example "Cz-M" refers to the montage which is the Cz channel referenced to the contralateral mastoid channel.

In addition, we created one manually selected montage for each subject (denoted "CI-closest" montage) with the active electrode positioned as close as possible to the subject's cochlear implant sound processor which was not blocked by the external coil. The aim was to collect data from the scalp that would be somewhat similar to that which would be collected by using implanted electrodes of the cochlear implant. T7 or T8 was selected as the active channel in 18 cases for left and right implants respectively (P7 and C6 were used instead for subjects 9 and 12 respectively because of coil placement). The reference channel was selected subjectively to be the closest to the extra-cochlear plate electrode of the Nucleus cochlear implant. Many of these channels were disconnected due to the coil placement. In those cases, the mastoid ipsilateral to the implant was selected as the reference channel (13 cases). Fig. 1 Panel A shows an example of subject 18's Cz-M and CI-closest montages, as well as the two disconnected channels electrodes P8 and P10 (stacked on cochlear implant cartoon), and two noisy channels TP8 and P2 (red crosses). Following montage determination, all montages of right sided implant subjects were mirrored horizontally in the analysis of the data; thus, all data were analysed as if the implant was on the left side.

Pre-processing was performed in MATLAB (version 2016b, MathWorks, MA, USA). Continuous data were zero-phase filtered between 1 and 45 Hz with elliptical filters designed with a passband ripple of < 1% and stopband attenuation of > 40 dB. Data were then down-sampled to 256 Hz and epoched in a window -600 ms to +1200 ms relative to stimulus onset. Any epochs with amplitudes that exceeded ±100μV or RMS voltage that exceeded 300 μV^2 were then rejected as artefact contaminated epochs. At least 250 artefact-free responses (corresponding to approximately 6 minutes of recording) were acquired at each stimulus level for each subject, and the collection of these responses are hereby referred to as blocks. After pre-processing, the dataset contained blocks of epochs from 20 subjects, each with six stimulus intensities. In the analysis, we used a subset of the epochs to examine the effect of recording time versus threshold estimation accuracy.

2.4 Feature extraction and growth function fitting

Two features were individually extracted from the data recorded at each montage. The first feature, peak-to-peak amplitude, was extracted by taking the average of each block of epochs and then taking the peak-to-peak voltage potential in a 50 to 500 ms post-stimulus window. The second feature, peak phase-locking value (peak PLV), was calculated by first taking a spectrogram of each epoch in a block with a 400 ms Hamming window and 20 ms time steps, and then calculating the PLV at each time-frequency point according to the following formula, where N is the number of epochs and θ represents the phase angle at each time (t) and frequency (f) point.

$$\mathrm{PLV}(t,f)=\frac{1}{N}\sqrt{\left[\sum_{i=1}^{N}\cos\left(\theta_i(t,f)\right)\right]^2+\left[\sum_{i=1}^{N}\sin\left(\theta_i(t,f)\right)\right]^2}$$

The peak PLV was extracted from the PLV spectrogram by taking the maximum PLV in a 1 to 20 Hz and 50 to 600 ms window. The PLV is in dimensionless units, and can range between zero (no similarity between epochs) and one (every epoch has exactly the same phase at a particular time-frequency point).

Each feature was bootstrapped for 100 iterations by sampling epochs with replacement and the median of the bootstrap distribution was used in growth function fitting. Exponential growth functions of the following form for peak-to-peak amplitude or PLV features were fitted to the features extracted from data for 10, 20, 40, 60 and 100% DR stimulus levels:

$$f(x) = a(1 - e^{-\frac{x-b}{c}})$$

Where x is the % DR, a is the upper asymptote of the feature growth function, b is the x-axis intercept and c controls the rate of growth of the exponential function.

2.5 Threshold level estimation

Fig. 1 Panel B demonstrates the proposed threshold estimation procedure using an example from subject 11. Following the growth function fitting, the sub-threshold response’s feature, in this case its phase locking value, was used to calculate the feature baseline level. The growth function was used to estimate the threshold level by finding the % DR on the x-axis where the function was equal to the baseline phase-locking value. Objective threshold estimation quality was quantified by comparing goodness of fit *r* values between the behavioral thresholds and objectively determined threshold estimates, as well as the standard deviation of the differences between the two.

<<Insert Figure 1 about here, in color>>

3 Results

Fig. 2 shows the measured behavioral threshold and comfort levels for each subject.

<<Insert Figure 2 about here, in color>>

3.1 Comparison of montage position and feature type

Fig. 3 Panels A to D show representative data from subject 18's response to stimuli presented at 100% DR. Features from the Cz-M montage (Panels A and B), which tends to have the most reliable responses (Abbas and Brown, 2015), can be compared to those from the CI-closest montage (Panels C and D). It can be observed that both peak-to-peak amplitude and peak phase-locking value features were smaller when using the CI-closest montage. Additionally, the CAEP was observed to be different in the two montages, visually evident both in the time domain (Panels A and C) and the phase spectra (Panels B and D). No stimulus artefact was visually present in the recording.

Panel E shows the ratio between feature amplitudes from the two different montages in all 20 subjects at each stimulation level (10, 20, 40, 60 and 100% DR). In response to 100% DR stimuli for example, the peak-to-peak amplitude feature calculated from the CI-closest montage was on average 52.6% of the size of the same feature calculated from the Cz-M montage, while for the peak PLV feature the ratio was 72.0%. To compare the feature amplitude ratio from the different recording locations, a general linear model was used with subjects as a random factor, and stimulus level and feature as fixed factors, as well as a level-feature interaction term. We found a significant effect of stimulus level ($F_{4,171} = 13.77$, $p <$ 0.001): the ratio was generally decreasing with increasing level. We also found a significant effect of feature ($F_{1,171}= 14.62$, $p < 0.001$): the ratio was smaller for peak-to-peak amplitude.

There was no significant interaction between level and feature ($F_{4,171} = 0.73$, $p = 0.573$); although an interaction effect may have been obscured by the large inter-subject variation. Finally, we observed that the ratio is above 1 for many subjects at the lower stimulation levels. This was likely due to noise levels dominating the recordings at each location.

This result suggests that the peak phase-locking value measure may be more resistant to changes in the recording montage than the peak-to-peak amplitude feature. However, this does not directly imply that the growth function fit and threshold estimation was better using the peak PLV. In the following analyses, we focused on the peak phase-locking value feature for threshold estimation and used the peak-to-peak amplitude as a comparison.

<<Insert Figure 3 about here, in color>>

3.2 Using PLV growth function to estimate hearing threshold in CI users

Estimated thresholds were determined for each subject using the method described in Fig. 1B. Each circle on Fig. 4 corresponds to one subject, with behaviorally-determined threshold on the x-axis, and the threshold estimate derived from the PLV growth functions on the y-axis. Thresholds estimated from data recorded from Cz-M are shown in Fig. 4A, and from the CI-closest montages in Fig. 4B. Both montages produced objective estimates of threshold that were highly correlated with the behavioral thresholds (Cz-M: $r = 0.979$, $p < 0.001$; CI-closest: $r = 0.966$, $p < 0.001$). To further assess the reliability of the objective threshold estimates, they were first converted to percent behavioral dynamic range (% DR). The standard deviation (SD) represents how the threshold estimate varies across subjects, and therefore how reliable the estimation method is. For the Cz-M montage, the across subjects SD was 5.6% DR (mean +9.3% DR, range –5.6 to +18.4% DR). For the CI-closest montage, the across subjects SD was 7.2% DR (mean +10.6, range -4.2 to +24.3% DR). A two-sample *F*-test on the threshold estimates in units of % DR revealed no significant difference between recording from the Cz-M and the CI-closest montages ($F_{19,19} = 0.595$, $p = 0.265$).

<<Insert Figure 4 about here, in color>>

The threshold estimates generated by the peak PLV feature were compared against the estimates obtained using peak-to-peak amplitude feature, at both recording locations. This analysis was performed by comparing the standard deviations of threshold estimates across subjects for both features at both recording locations, a total of four conditions. Fig. 5 shows the standard deviations and their 95% multiple comparison confidence intervals, as generated in Minitab (version 18.1, Minitab Inc., PA, USA). The multiple comparisons test showed a significant difference in standard deviation ($p = 0.047$), and post-hoc analysis showed that there was only a significant difference between peak-to-peak amplitude using the two different recording locations. This is evidence that the peak PLV feature is more robust to recording location changes than the peak-to-peak amplitude feature.

<<Insert Figure 5 about here, in grayscale>>

To further examine the other placement locations' threshold estimation performance, we compared the performance of Cz-M data against each of the other left-implant-normalized scalp channels while using the peak phase-locking value feature. P7-M and TP7-M were excluded from analysis as more than half were disconnected due to obstruction by the subjects' implant coils. We bootstrapped the *r* values (100 iterations) for each of the remaining 62 channels and found that no montage performed significantly better than Cz-M (empirical distribution comparison using $p < 0.05$), although FPz-M performed the best ($r = 0.990$, compared to Cz-M where $r = 0.979$). Only six channels (PO7, TP8, P6, O1, Oz, O2) performed significantly worse. These were all occipital channels or channels near the implant-contralateral mastoid that created montages which may have not captured the cortical response well for growth function fitting. These results show that recording montage placement precision is not important for the best growth function threshold estimate; rather, the care should be taken to not place the active electrode too close to the reference electrode or the occipital region of the scalp when using the growth function method.

3.3 The effect of reducing number of epochs on the threshold estimate

To further investigate clinical viability, the effect of reducing the number of epochs used for the threshold estimate was calculated. Data recorded from the Cz-M and CI-closest montages were used, and threshold estimation accuracy was compared while limiting the data to between 20 and 200 epochs per stimulus intensity in steps of ten epochs. The number of epochs used was converted into a 'testing time' value by multiplying by the average inter-stimulus interval (1.5 seconds). In the growth function fitting, the -50% DR response data (a stimulus level very unlikely to be perceived) was substituted with background EEG activity, thus reducing the number of stimulus levels needed from six to five. The threshold estimation procedure (as described in Fig. 1B) was performed in each subject 50 times, each time iterating randomly selected epochs within each stimulus block, for a total of 1000 simulated runs per ‘testing time’. Results where the baseline PLV feature was greater than the asymptote of the fitted growth function or where the threshold estimate was less than 0 or greater than 255 current levels were rejected as invalid fits.

Fig. 6 Panels A and B show histograms of threshold estimation accuracy after a simulated test time of five minutes and ten minutes to test 5 stimulus levels, corresponding to 40 and 80 epochs per level respectively. Fig. 6 Panels C and D show the results for all recording times simulated in terms of *r* value and estimate difference in % DR, respectively. In Panel D, a wider spread of threshold estimates calculated from data recorded at the CI-closest montage can be seen compared to the Cz-M montage, implying a lower reliability of estimation. Table 3 summarizes these results. At five minutes test time, there was a significant difference between the standard deviation of the estimates calculated from data recorded at the two montages (two sample *F*-test, $F_{967,851} = 0.652$, $p < 0.0001$). The significant difference remains at ten minutes test time ($F_{990,950} = 0.338$, $p < 0.0001$). These results show that data from the CI-closest montage do not produce threshold estimates that are as reliable as data from the Cz-M montage. However more testing time can compensate for the less accurate estimate (Fig. 6C). It is possible that the non-significant difference between the two recording locations when using all data (seen in previous section and in Fig. 4) is due to the asymptotic effect on the threshold estimate accuracy when more data are used. These results indicate that lower threshold estimation accuracy obtained from the CI-closest montage can be compensated for by recording more data. Lastly, the same analysis was performed with the peak-to-peak amplitude feature applied to the CI-closest location as a comparison. We found that the peak PLV performed better by having a lower standard deviation of threshold estimate than the peak-to-peak amplitude at 5 minutes testing time ($F_{851,825} = 0.651$, $p < 0.0001$) and at 10 minutes testing time ($F_{950,933} = 0.508$, $p < 0.0001$).

<<Insert Figure 6 about here, in color>>

Table 3. Summarized results from Fig. 6 Panels A and B. Threshold estimation accuracy for the two recording locations and two testing times are expressed in correlation coefficient and % DR deviation from behavioral threshold.

Testing time	**Montage**	**Correlation coefficient**	**Difference between threshold estimate and behavioral threshold**	**Number of invalid fits (out of 1000 simulated runs)**

			across subjects (mean ± 1 std.) (% DR)	
5 minutes, equivalent to 5 blocks of 40 epochs	Cz-M	0.823	12.9 ± 19.9	32
	CI-closest	0.723	14.2 ± 24.7	148
10 minutes, equivalent to 5 blocks of 80 epochs	Cz-M	0.924	11.1 ± 11.1	9
	CI-closest	0.808	13.5 ± 19.1	49

3.4 Comfort level estimation using population and growth function information

The data were further examined to see if the PLV growth functions could be used to determine comfort level. A trial method to estimate comfort levels based on population values of PLV was developed, and then evaluated with a leave-one-out cross validation.

Fig. 7 illustrates the proposed comfort level estimation method. The procedure assumed that the comfort level corresponds to a feature amplitude that is at a fixed proportion, *m*, of the amplitude of the feature growth function asymptote:

$$m = \frac{PLV_C - PLV_b}{PLV_{asym} - PLV_b}$$

Here, PLV_C is the peak PLV at the behavioral comfort level, PLV_{asym} is the asymptotic PLV based on the growth function fit, and PLV_b is the base PLV calculated from the sub-threshold level. The procedure involves determining *m* from a population with known comfort levels (Fig. 7A) and applying this population value of *m* to an individual with unknown comfort level to estimate their comfort level (Fig. 7B). Fig. 7A uses subject 11 as an example of calculating *m* for an implant user whose behavioral comfort level is known. This performed for all subjects except 18, and the population-averaged *m* is denoted m_{pop}. Fig. 7B illustrates the leave-one-out method for determining comfort level for subject 18, Given m_{pop}, we find the comfort level estimate as such:

$$PLV_C\ estimate = m_{pop}(PLV_{asym} - PLV_b) + PLV_b$$

Finally, we use subject 18's growth function to predict the stimulus intensity that corresponds to their estimated PLV_C.

<<Insert Figure 7 about here, in color>>

We found a moderate correlation between comfort level estimates calculated with the data acquired from the Cz-M montage and the behavioral comfort level (r = 0.50, p = 0.024, see Fig. 8A). However, there was no significant correlation when using data acquired from the CI-closest montage (r = 0.14, p = 0.565, see Fig. 8B). The comfort level estimation result is not sufficiently accurate for clinical application.

<<Insert Figure 8 about here, in color>>

4 Discussion

In this study, a fully objective method for estimating the hearing threshold for a cochlear implant electrode was developed using single-channel EEG recordings combined with phase-locking values and growth function fitting. The method can generate objective threshold estimates that correlate well with behavioral thresholds. This method can therefore be used for improving mapping procedures in adult and potentially infant implant recipients. Further, we have shown that data recorded from a location physically close to the cochlear implant can generate equally accurate threshold estimates, albeit with longer recording times. Future cochlear implant systems which incorporate recording electrodes either implanted or mounted on the external sound processor may benefit from using phase-based objective measures such as the peak phase-locking value feature investigated in this study.

4.1 Additional information from the growth functions

An interesting observation was made from Fig 6D: on average, threshold estimates were overestimated by around 10% DR compared to the behavioral threshold. More specifically, Fig. 4 showed this overestimate was higher when the behavioral thresholds were lower. While the mechanisms that contribute to these observations are unclear, it is possible that the lower behavioral thresholds correspond to better neural survival in the cochlea. Better neural survival may mean more stochastic responses across epochs due to an abundancy of available neurons to respond to the sound, leading to a lower phase-locking value. A lower-valued PLV growth function, along with the assumption that the baseline phase-locking value is unchanged, would lead to an overestimate of the behavioral hearing threshold. Along with the finding that threshold estimates using CAEP global field power growth functions have greater uncertainty when behavioral thresholds are lower (Visram et al., 2015), further exploration of the interaction of the growth function characteristics with behavioral threshold levels is needed.

In our previous work (Mao et al., 2018), we used the same hearing threshold estimation method but in acoustic listeners. The growth functions that were produced in the two studies are difficult to directly compare for various reasons: firstly, there was a large age difference (previous study: mean = 25 years; this study: mean = 65 years); secondly, the range of stimulus intensities used is hard to compare as the levels were of different units (percent dynamic range and dB SL); finally, comfortable level stimuli were not analysed in the previous study, so the range of loudness of the stimuli differed. Regardless, a comparison could be made by observing the PLV growth function asymptotes. The acoustic listeners reached an asymptotic PLV of 0.86 ± 0.17 (mean ± 1 std. across 20 subjects), whereas the cochlear implant listeners reached a PLV asymptote of 0.49 ± 0.20 (mean ± 1 std. across 20 subjects). The significantly lower PLV asymptote (Two-sample unpaired t-test, $t = 3.43$, $p = 0.0015$) for cochlear implant listeners may be due to two factors. First, the growth function asymptotes for acoustic listeners may have large uncertainties. Since the acoustic listeners were only tested to 60 dB SL, it was possible that their growth functions did not approach an asymptotic PLV value; rather, the PLV values may have been in the linear region of feature growth. To compare, we extracted the 95% confidence interval for the asymptote during curve fitting. We found that the 95% confidence interval for the fitted asymptote was 4.15 in the acoustic listeners compared to 0.93 in the cochlear implant listeners (significantly higher; rank-sum test: $p = 0.001$). An unreliable estimate for asymptote, however, did not affect the threshold estimation accuracy (Mao et al., 2018). Second, the older age of the cochlear implant users may contribute, although this has not been shown in literature; it is only known that CAEP magnitude decreases with age (Goodin et al., 1978). In our study with 20 subjects, there was no correlation between growth function asymptote and age (Pearson correlation, $r = -0.25$, $p = 0.29$).

If the asymptote comparisons are reliable and there is lower maximal phase-locking for cochlear implant listeners, it may mean that individual responses to the same stimuli are more varied. This may have implications on speech understanding; individuals may have more trouble distinguishing speech tokens if, for example, their perception of the same speech token is already different each time. Thus, an interesting experiment would be to test whether maximal phase-locking values were correlated with speech understanding.

4.2 Comfort level estimation

There are currently no cortical response features described in literature that correlate well with comfort levels of hearing. The use of PLV growth functions as tried with the current data did produce a significant correlation, but with insufficient accuracy to be clinically useful. The comfort level estimates varied greatly in accuracy across subjects, and this variation may have at least two sources. First, there is variation in the behavioral estimate due to the subjectivity of estimates given by the subjects, and there is no way of knowing how consistent “comfortably loud” is across subjects. Second, there is no evidence that the absolute phase-locking value can reflect loudness perceived. This is clear when comparing the much lower levels of PLV in the cochlear implant group compared to the normal hearing group in the previous study, despite the loudness range being greater in the cochlear implant group. The algorithm attempted to account partially for this by comparing ratios of asymptotic peak phase-locking with comfort level peak phase-locking instead of directly comparing absolute phase-locking values across subjects; however, the results were still not sufficiently accurate to be clinically useful. This is also initial evidence that the phase-locking value feature is a poor measure of loudness.

4.3 Towards a fully objective clinical tool for cochlear implant mapping

There are still challenges remaining to the implementation of the totally objective threshold estimation technique in clinical settings. Some device manufacturers do not use the threshold level in their fitting procedure (for example, Advanced Bionics implants are fitted at the comfort level only), thus this method may not be applicable to these devices. Another limitation of all current methods that utilize cortical responses is the need to use scalp electrodes and additional measurement equipment. Unfortunately, the neural telemetry systems that exist in current implants to record ECAPs from intracochlear electrodes can only record data for several milliseconds at a time, and therefore cannot fully capture a cortical response, which is hundreds of milliseconds in duration. There are two possible workarounds:

adapt existing hardware by concatenating multiple windows of responses (Campbell et al., 2015; Mc Laughlin et al., 2012), or develop improved or additional recording hardware in cochlear implants. In either case, measuring the cortical response using electrodes in the cochlear implant system currently requires that the recording electrode positions are distant from the ideal location of Cz. A recording location near the implant gives rise to two issues: larger stimulation artefacts (Mc Laughlin et al., 2013) and reduced response magnitudes (Kelly et al., 2005).

In this study, the stimulation artefacts were removed by using a short (50 ms) stimulus pulse train, so that the response of interest (after 50 ms from stimulus onset) was not affected. Clinically, a 500-ms duration pulse-train is used to find behavioral threshold. Stimulus duration may affect perception differently in different individuals; thus, the thresholds estimate from the cortical response may be less correlated to the 500-ms behavioral thresholds than was found in this study for the 50-ms behavioral thresholds. One possible solution to account for the threshold difference is to use a correction factor to account for the average difference between 50-ms and 500-ms thresholds, which have been found to be correlated (Visram et al., 2015). Another solution is to use a 500-ms duration stimulus and perform stimulation artefact correction with a single channel method (Mc Laughlin et al., 2013). However, it is uncertain whether the artefacts can be completely removed with that method (Mc Laughlin et al., 2013). Any leftover stimulus artefact can affect the growth function fitting greatly, as they appear systematically in each response block and will artificially elevate the fitted growth function. It is likely that, for the estimation method presented in the current work, correcting for a shorter stimulus duration may be a more suitable approach than risking residual stimulation artefact by using a longer stimulus.

The current results showed that measuring cortical responses from electrodes located near the implant itself will lead to reduced PLV values compared to measurement positions near Cz. Further, the lower PLV magnitude lengthens the recording time required to reach the same accuracy and reliability as obtained from data collected at Cz. Thus, the study supports the potential for the use of the cochlear implant to perform recording via reverse telemetry systems. As mentioned above, this can be achieved with either an improved recording capability of the existing neural telemetry systems, or the implementation of a behind-the-ear recording electrode on the speech processor itself. An improved neural telemetry system would use an intra-cochlear electrode and extracochlear electrode as the active and the reference electrodes. Thus, the orientation of the pair of recording electrodes may be slightly different to what we used in the current study (mainly T7/T8 to the ipsilateral mastoid), and our result may not be indicative of the performance when using two implanted electrodes. Instead, our result is more predictive of performance when a behind-the-ear electrode is referenced to an extra-cochlear electrode, for example.

We only tested a single, mid-array stimulation electrode in this study. Generally, stimuli on the more basal electrodes (corresponding to higher frequencies) generate lower peak-to-peak cortical response amplitudes (Firszt et al., 2002), causing possible poorer growth function fitting and thus worse threshold estimates. Visram et al. (2015) found that reliable growth functions of global field power could not be fitted for three out of nine subjects for a basal electrode (electrode 3), whereas the function for a mid-array electrode could be fitted for all subjects. Further work is required to determine whether our proposed single-channel, phase-locking feature approach is applicable along the whole implant array. It is possible that the phase-locking value would be less affected than other features by smaller response amplitudes, as suggested by the results of this study, and thus might generate accurate threshold estimates for all electrodes.

The preparation of subjects took approximately 15 minutes because a full scalp EEG cap was used in this study. In a clinical setting, only a small subset of scalp electrodes would be needed (a minimum of three would be needed to perform the analysis presented in this paper) which may take two to three minutes. Thus, application of this method in clinic is as straightforward as testing an ABR. Further, if the recording was incorporated into the implant, the preparation time would be eliminated entirely.

5 Conclusion

In this study, a fully objective and reliable method for hearing threshold estimation in cochlear implant users was demonstrated, using the cortical auditory response, the phase-locking value and a single recording channel. The PLV method was robust to differences in scalp recording locations, , provided additional measurement time. A future goal is the implementation of a telemetry system with measurement electrodes built into an implant or behind-the-hear device, which would be able to record cortical responses and thus perform fully automated threshold estimation. A more accurate comfort level estimation method is needed for fully objective implant fitting.

6 Acknowledgements

Darren Mao was supported by an Australian Government Research Training Program Scholarship. Colette McKay was supported by a Veski innovation fellowship. Hamish Innes-Brown was supported by a National Health and Medical Research Council (Australia) Early-Career Research Fellowship (Grant number 1069999). The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program.

7 References

Abbas, P.J., Brown, C.J. 2015. Assessment of responses to cochlear implant stimulation at different levels of the auditory pathway. Hear Res 322, 67-76.

Brown, C.J., Lopez, S.M., Hughes, M.L., Abbas, P.J. 1999. Relationship between EABR thresholds and levels used to program the CLARION® speech processor. Ann Otol Rhinol Laryngol 108, 50-57.

Brown, C.J., Hughes, M.L., Luk, B., Abbas, P.J., Wolaver, A., Gervais, J. 2000. The relationship between EAP and EABR thresholds and levels used to program the nucleus 24 speech processor: Data from adults. Ear Hear 21, 151-163.

Campbell, L., Kaicer, A., Briggs, R., O'Leary, S. 2015. Cochlear response telemetry: intracochlear electrocochleography via cochlear implant neural response telemetry pilot study results. Otol Neurotol 36, 399-405.

Cone, B., Whitaker, R. 2013. Dynamics of infant cortical auditory evoked potentials (CAEPs) for tone and speech tokens. International journal of pediatric otorhinolaryngology 77, 1162-1173.

Delorme, A., Makeig, S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134, 9-21.

Donaldson, G.S., Viemeister, N.F., Nelson, D.A. 1997. Psychometric functions and temporal integration in electric hearing. J Acoust Soc Am 101, 3706-21.

Firszt, J.B., Chambers, R.D., Kraus, Reeder, R.M. 2002. Neurophysiology of cochlear implant users I: effects of stimulus current level and electrode site on the electrical ABR, MLR, and N1-P2 response. Ear Hear 23, 502-15.

Goodin, D.S., Squires, K.C., Henderson, B.H., Starr, A. 1978. Age-related variations in evoked potentials to auditory stimuli in normal human subjects. Electroencephalogr Clin Neurophysiol 44, 447-58.

Hodges, A.V., Ruth, R.A., Lambert, P.R., Balkany, T.J. 1994. Electric auditory brain-stem responses in nucleus multichannel cochlear implant users. Arch Otolaryngol Head Neck Surg 120, 1093-9.

Hofmann, M., Wouters, J. 2010. Electrically evoked auditory steady state responses in cochlear implant users. J Assoc Res Otolaryngol 11, 267-82.

Hughes, M.L. 2013. Objective measures in cochlear implants San Diego : Plural Pub., c2013.

Kelly, A.S., Purdy, S.C., Thorne, P.R. 2005. Electrophysiological and speech perception measures of auditory processing in experienced adult cochlear implant users. Clin Neurophysiol 116, 1235-46.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., Naatane, R. 2002. Maturation of the auditory event-related potentials during the first year of life. Neuroreport 13, 47-51.

Mao, D., Innes-Brown, H., Petoe, M.A., Wong, Y.T., McKay, C.M. 2018. Cortical auditory evoked potential time-frequency growth functions for fully objective hearing threshold estimation. Hearing Research.

Mc Laughlin, M., Lu, T., Dimitrijevic, A., Zeng, F.G. 2012. Towards a closed-loop cochlear implant system: application of embedded monitoring of peripheral and central neural activity. IEEE Trans Neural Syst Rehabil Eng 20, 443-54.
Mc Laughlin, M., Lopez Valdes, A., Reilly, R.B., Zeng, F.G. 2013. Cochlear implant artifact attenuation in late auditory evoked potentials: a single channel approach. Hear Res 302, 84-95.
McKay, C.M., Smale, N. 2017. The relation between ECAP measurements and the effect of rate on behavioral thresholds in cochlear implant users. Hear Res 346, 62-70.
McKay, C.M., Chandan, K., Akhoun, I., Siciliano, C., Kluk, K. 2013. Can ECAP Measures Be Used for Totally Objective Programming of Cochlear Implants? J Assoc Res Otolaryngol 14, 879-890.
Miller, C.A., Brown, C.J., Abbas, P.J., Chi, S.L. 2008. The clinical application of potentials evoked from the peripheral auditory system. Hear Res 242, 184-97.
Näätänen, R., Picton, T. 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375-425.
Shapiro, W.H. 2014. Advancements in Cochlear Implant Programming. In: Waltzman, S.B., Roland, J.T., (Eds.), Cochlear Implants, 3rd. ed. Thieme Medical Publishers, New York. pp. 148-157.
Silva, L.A., Couto, M.I., Matas, C.G., Carvalho, A.C. 2013. Long latency auditory evoked potentials in children with cochlear implants: systematic review. Codas 25, 595-600.
Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J. 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 16, 4240-9.
Vaerenberg, B., Smits, C., De Ceulaer, G., Zir, E., Harman, S., Jaspers, N., Tam, Y., Dillon, M., Wesarg, T., Martin-Bonniot, D., Gartner, L., Cozma, S., Kosaner, J., Prentiss, S., Sasidharan, P., Briaire, J.J., Bradley, J., Debruyne, J., Hollow, R., Patadia, R., Mens, L., Veekmans, K., Greisiger, R., Harboun-Cohen, E., Borel, S., Tavora-Vieira, D., Mancini, P., Cullington, H., Ng, A.H., Walkowiak, A., Shapiro, W.H., Govaerts, P.J. 2014. Cochlear implant programming: a global survey on the state of the art. ScientificWorldJournal 2014, 501738.
Visram, A.S., Innes-Brown, H., El-Deredy, W., McKay, C.M. 2015. Cortical auditory evoked potentials as an objective measure of behavioral thresholds in cochlear implant users. Hear Res 327, 35-42.
Wunderlich, J.L., Cone-Wesson, B.K., Shepherd, R. 2006. Maturation of the cortical auditory evoked potential in infants and young children. Hear Res 212, 185-202.