Research paper

Assessing temporal modulation sensitivity using electrically evoked auditory steady state responses

Robert Luke[*], Lieselot Van Deun, Michael Hofmann, Astrid van Wieringen, Jan Wouters

ExpORL, Department Neurosciences, KULeuven, O&N 2, Herestraat 49 bus 721, 3000 Leuven, Belgium

ABSTRACT

Temporal cues are important for cochlear implant (CI) users when listening to speech. Users with greater sensitivity to temporal modulations show better speech recognition and modifications to stimulation parameters based on modulation sensitivity have resulted in improved speech understanding. Behavioural measures of temporal sensitivity require cooperative participants and a large amount of time. These limitations have motivated the desire for an objective measure with which to appraise temporal sensitivity for CI users.

Electrically evoked auditory steady state responses (EASSRs) are neural responses to periodic electrical stimulation that have been used to predict threshold (T) levels. In this study we evaluate the use of EASSRs as a tool for assessing temporal modulation sensitivity.

Modulation sensitivity was assessed behaviourally using modulation detection thresholds (MDTs) for a 20 Hz rate. On the same stimulation sites, EASSRs were measured using sinusoidally amplitude modulated pulse trains at 4 and 40 Hz. Measurements were taken using a bipolar configuration on 12 electrode pairs over 5 participants. Results showed that EASSR amplitudes and signal-to-noise ratios (SNRs) were significantly related to the MDTs. Larger EASSRs corresponded with sites of improved modulation sensitivity. This relation was driven by across-subject variation. This result indicates that EASSRs may be used as an objective measure of site-specific temporal sensitivity for CI users.

1. Introduction

Cochlear implant (CI) recipients often understand speech well in quiet conditions but in difficult listening environments their performance worsens and becomes variable. Pre-, per- and post-operative factors account for 22% of this variance (Lazard et al., 2012). A proposed cause for some of the remaining variability in performance is perceptual variance along the tonotopic axis caused by the quality of each electrode neuron interface (ENI) (Pfingst et al., 2008; Bierer and Faulkner, 2010). Reducing these perceptual variations, by adjusting the stimulation parameters of individual sites, has been suggested as a means for improving speech performance (Zwolan et al., 1997; Pfingst et al., 2008).

The ENI affects the ability of an implanted electrode to transmit information to the auditory nerve. Ideally the electrode lies close to the modiolus, which should contain a full compliment of spiral ganglion cells (SGCs) (Long et al., 2014). Variations in electrode placement, tissue growth and local degeneration of SGCs, will cause variations in the ENI and differences in the perception of both spectral and temporal cues.

To account for individual variation along the implanted array, every device is fitted by an audiologist. The fitted parameters for each electrode include the threshold level (T) and comfort level (C). The parameters are stored in the device, and referred to as the MAP.

Commercial CIs transmit both spectral and temporal cues (Xu et al., 2005). Spectral information is predominantly transmitted through the location of stimulated electrodes, and is distorted by current spread and loss of SGCs. Using a focused tripolar mode, stimulation sites with high T levels have been related to broad psychophysical tuning curves which may indicate dead regions in electrical hearing (Bierer and Faulkner, 2010). High variability in T levels across electrodes negatively affects speech performance (Pfingst et al., 2004; Bierer, 2007; Long et al., 2014), possibly due to distortion of the internal representation of the spectrum of the signal.

Compared to normal hearing listeners, CI recipients have reduced access to spectral cues (Friesen et al., 2001). This places increased importance on temporal sensitivity, which is commonly

* Corresponding author. Tel.: +32 16 330475.
  E-mail address: robert.luke@med.kuleuven.be (R. Luke).

assessed using modulation detection thresholds (MDTs). MDTs of CI users have been related to consonant and vowel recognition (Fu, 2002) and to word recognition and speech reception thresholds (SRTs) (Won et al., 2011).

The quality of the ENI varies uniquely along each implanted array. Altering the MAP based on the performance of each ENI has resulted in improved speech performance. Site specific adjustments have been made using a variety of selection criteria and adjustment methods. Zwolan et al. (1997) used 200 ms pulse trains to determine which channels along the electrode array could be discriminated from each other. Channels that were indiscriminable from each other were deactivated. With this altered MAP, seven of nine subjects improved in at least one speech recognition measure. Garadat et al. (2012) used masked MDT performance to create two 10 channel MAPs, one with good across-site mean MDT performance and one with poor performance. The MDTs were determined in the presence of an interleaved masker on the adjacent apical site. The array was divided into five sections of four electrodes. In each MAP two electrodes were retained from each section. One MAP retained two electrodes per section that exhibited the best masked MDTs and the second MAP retained the other two electrodes per section. MAPs with better across-site mean MDTs resulted in better speech recognition. Garadat et al. (2013) extended the previous study by creating a MAP for each participant that improved the mean modulation sensitivity while only removing five electrodes. They endeavoured to remove sites in a distributed fashion across the array, but did not always remove electrodes from all regions of the array. The frequency allocation was redistributed across the remaining electrodes. The modified MAP resulted in a mean SRT improvement of 2 dB over the clinical map and led to better performance than the clinical map for consonant recognition but not for vowel recognition. Zhou and Pfingst (2014) increased the T level of the five electrodes with the poorest MDTs. The T level was increased to artificially increase the loudness of the channel, which improves modulation sensitivity. This adjustment resulted in a mean SRT improvement of 2.4 dB.

Psychophysical evaluation of stimulation sites has illustrated the potential benefit of site-specific adjustments, but these behavioural measures are not always clinically feasible due to their extensive testing time and need for a cooperative participant. Objective measures based on evoked potentials offer the possibility of fast automated evaluation of stimulation sites. Electrically evoked auditory brain stem responses (EABRs) have been used to predict high thresholds and thus sites with poor spectral sensitivity (Bierer et al., 2011; Brown et al., 1990). But neither EABRs nor electrically evoked compound action potentials have shown clinically useful correlations with speech perception tasks or temporal sensitivity (Miller et al., 2008). Here we propose electrically evoked auditory steady state responses (EASSRs) as a measure of site specific temporal sensitivity.

Auditory steady state responses (ASSRs) are neural responses to periodic auditory stimuli (Galambos et al., 1981; Picton et al., 2003) that can be used to predict frequency specific behavioural hearing thresholds (Rance et al., 1995). ASSRs have been related to phoneme recognition, word recognition, word discrimination and speech in noise perception (Dimitrijevic et al., 2001; Picton et al., 2001; Dimitrijevic et al., 2004; Alaerts et al., 2009; Poelmans et al., 2012).

EASSRs can be measured for CI recipients. These recordings are distorted by artifacts from radio frequency transmission and electrical stimulation. Removal of these artifacts (Hofmann and Wouters, 2010) has allowed prediction of behavioural thresholds at clinically relevant pulse rates (Hofmann and Wouters, 2012). We hypothesise that stimulation sites with increased neural responses to modulated auditory input will correspond to sites with improved modulation sensitivity. Thus EASSRs will provide an objective method for assessing the temporal sensitivity of cochlear implant stimulation sites.

2. Materials and methods

2.1. Participants

Five native Flemish-speaking participants volunteered for this experiment. All participants were CI patients of the ENT Department at the UZ Leuven University Hospitals. The details of each participant are included in Table 1, including their word recognition in sentences as evaluated using the LIST sentences (Van Wieringen and Wouters, 2008). Testing was approved by the Medical Ethics Committee of the UZ Leuven (approval number B32220072126) and informed consent was obtained.

2.2. Experiment

Each participant took part in four sessions, each lasting between two and four hours. The measures in successive sessions were: (1) T and C level measurements and loudness balancing, (2) EASSR measurements, (3) loudness balancing for the modulation detection task and (4) modulation detection task. T and C levels were checked again in sessions 2–4.

All stimuli consisted of symmetric biphasic pulse trains with 60 μs phase width and 8 μs inter phase gap, presented at a rate of 900 pulses per second in bipolar (BP) mode (for further information on the choice of stimulation parameters, see Section 4). All stimuli were delivered using the Cochlear Nucleus Implant Communicator (NIC). Bipolar stimulation was used as it may stimulate a more localised region of the cochlea than for monopolar stimulation (Snyder et al., 2008; Kwon & van den Honert, 2006). Pulse polarity is described relative to the more apical electrode. Cathodic first stimulation is defined as the biphasic negative phase first. All psychophysical modulation detection tasks were conducted using cathodic first stimulation. EASSR recordings were obtained using both polarity configurations. This was done so that a comparison could be made to previously published EASSR results (Hofmann and Wouters, 2012). All stimulus magnitudes are reported in Cochlear clinical current units (cu), which is a logarithmic conversion from amperes. For the CIC3 implant, the conversion from cu to current is $i = 10 \times 10^{-6} \times 175^{cu/255} \mu A$, and for the CIC4 implant $i = 17.5 \times 10^{-6} \times 100^{cu/255} \mu A$.

Stimuli were presented on three electrode pairs for each participant. These electrode pairs were spaced along the array to excite basal, middle and apical regions along the cochlea. All participants used BP+2 mode except E9, who was unable to perceive any stimulus in this mode (Table 1). This participant had the mode changed to BP+5, for which T and C levels were reached on the basal and middle electrode pairs. On the apical electrode, participant E9 did not reach a comfortable percept and this site was excluded. The dynamic range (DR) is defined as the difference between C and T levels. Two electrode pairs were excluded (participants E1 and E2) because the local variation in T and C levels of the clinical monopolar MAP was more than half of the mean DR of these electrodes. The T levels of all included sites varied by <40% of mean DR. Both excluded sites were basal pairs and the T level varied by >69% of the mean DR. Greater differences were assumed to be a sign of highly varying ENI conditions, which cannot be unambiguously assessed with the currently available spatially wide stimulation patterns of bipolar stimuli.

Table 1
Participant identifier, age, years of deafness at implantation (for progressive loss, number of years with >80 dB HL), years of CI experience, type and side of implant, etiology, speech reception threshold as evaluated using the LIST sentences (dB SNR), mode of stimulation, electrodes used for EASSR and MDT experiments.

ID	Age	D	Exp	Implant	Side	Etiology	LIST	Mode	Basal	Middle	Apical
E1	52	22	2	CI24RE	R	Hereditary	1.3	BP+2		12−15	19−22
E2	63	6	2	CI24RE	R	Unknown	−2	BP+2		9−12	17−20
E5	66	5	13	CI24R	L	Noise induced	−1.5	BP+2	1−4	10−13	19−22
E7	76	?	7	CI24RE	R	Perinatal hypoxia	18.5	BP+2	1−4	11−14	19−22
E9	23	10	11	CI24R	L	Hereditary	20	BP+5	1−7	8−14	

2.3. Threshold, comfort & stimulation levels

Threshold levels for unmodulated stimuli (T_u) were determined for each stimulation site using an adaptive three-alternative forced-choice procedure. In each trial, three intervals of 500 ms separated by 500 ms of silence were presented. During each interval, a corresponding button on the computer screen changed color. One randomly selected interval contained an unmodulated pulse train with current that varied on each trial according to an adaptive two-down one-up rule (Levitt, 1971). No stimulus was presented in the other intervals. Participants were asked to identify the interval in which they heard a sound. After indicating which interval contained the stimulus, the next trial started. Six reversals were obtained for each run. The step size was 4 cu until the first two reversals occurred and was 2 cu thereafter. The staircase was visually checked for convergence and the cu values at the last four reversals were averaged and recorded as the threshold. This procedure was completed three times for each site, and results were averaged for each site to determine T_u.

The comfort level was defined as the maximum level the user judged they could listen to continuously for 3 h. Comfort levels (C_u) for unmodulated stimuli were determined by increasing the stimulation level from 0 cu until the participant indicated that the percept was uncomfortably loud. Then the level was decreased until the listener indicated that it was at the comfort level, which was recorded as. C_u

All stimuli were presented using a research speech processor (L34) provided by Cochlear Ltd. Apex3 was used to determine the unmodulated threshold and comfort levels (Francart et al., 2008).

2.4. Electrically evoked auditory steady state responses

2.4.1. Stimulus loudness balancing

The level of an unmodulated pulse train on the apical electrode was adjusted to be equally loud as stimulation at C_u on the middle electrode. This was done using an adjustment procedure consisting of four trials. Each trial continued until the participant indicated that the two sounds were equally loud. In the first trial the listener was repeatedly presented with the target, 1 s of unmodulated pulses at C_u on the middle electrode; silence for 500 ms; the variable stimulus, 1 s of unmodulated pulses on the apical electrode; and 1 s of silence. The listener was asked to adjust the variable stimulus until the two sounds were equally loud. If no adjustment was made the presentation continued at the same values. The variable stimulus was started below the expected balanced value. When the listener indicated that the loudness was equal, the difference in cu was recorded as the trial result. In the second trial, the target was changed to an unmodulated pulse train on the apical electrode at the level from the previous trial. The variable stimulus was an unmodulated pulse train on the middle electrode, with the initial value set below C_u. Again the listener had to adjust the variable stimulus until the loudness was equal. In the last two trials the target and variable stimuli were the same as for the first two trials, except that the variable stimulus was started above the expected

value. In all trials the participant was encouraged to continue past the point where they first found the loudness equivalent and then turn back if required. The average difference in the levels of the target and variable stimuli of the four trials was used to determine the current required for loudness balanced unmodulated pulse trains at C level ($C_{u,b}$). The same procedure was used to balance the basal electrode to the middle electrode.

The same four-trial procedure was used at each stimulation site to balance the loudness of a 40-Hz modulated signal to $C_{u,b}$. In the first trial of the procedure, the target was an unmodulated pulse train at $C_{u,b}$, and the variable stimulus was a sinusoidally modulated pulse train with current between T_u and $C_{m,b}$. The listener adjusted the value of $C_{m,b}$.

Loudness differences caused by stimulation polarity were removed by balancing opposing polarity configurations. All previous balancing was with cathodic-first stimuli. Modulated cathodic-first stimulation between T_u and $C_{m,b}$ was used as a target, and modulated anodic-first stimuli were balanced using the afore-described procedure.

All stimuli were presented via the L34. Loudness balancing was performed with custom scripts written in python interfacing with NIC.

2.4.2. Stimulation

EASSR stimuli were continuous amplitude-modulated pulse trains with modulation frequencies of 3.906 and 40.039 Hz, called for simplicity 4 and 40 Hz. 40 Hz was chosen, as modulation rates in this range give the largest ASSR SNRs (Picton et al., 2003). 4-Hz ASSRs produce smaller SNRs than those in the 40-Hz range, but can still be reliably detected (Alaerts et al., 2009). Both frequencies are within the range of perceptually relevant envelope modulation frequencies for CI users (Stone et al., 2008; Füllgrabe et al., 2009). Furthermore, 4 Hz was selected as it is a prominent frequency in the average envelope of natural speech (Edwards and Chang, 2013). Pulse amplitudes were sinusoidally amplitude modulated in amperes between T_u and $C_{m,b}$.

Each stimulus consisted of 256 epochs, each with a duration of 1.024 s. Stimuli were presented in both anodic- and cathodic-first configurations, resulting in four recordings per stimulation site. Stimuli were grouped by and randomised across frequencies, stimulation sites and polarities.

Participants were lying on a bed or sitting in a comfortable chair in a sound-proof and electromagnetically shielded room. Participants were asked to move as little as possible and to help them remain awake they watched a silent subtitled movie. Stimuli were generated using custom software as described in Hofmann and Wouters (2010) that interfaced with a L34.

2.4.3. Recording and postprocessing

The EEG was measured using a BioSemi 64-channel system with Ag/AgCl active scalp electrodes. Electrodes were placed according to the standard 10−20 position with the guidance of a headcap. Incoming signals were amplified and sampled at 8192 Hz. EEG data along with synchronisation signals were recorded using the

BioSemi actiview software, for offline processing in MATLAB. An artificial recording was generated for each modulation frequency that merged the two EASSR polarities by concatenating epochs of alternating polarity. This allowed for analysis of the two conditions jointly as well as independently of each other.

The EEG recorded during CI stimulation contains artifacts from the RF transmission and electrical stimulation. This can mask the neural responses and must be removed prior to further analysis. CI artifacts were removed using a method adapted from Hofmann and Wouters (2010), but this experiment used a different amplifier with a lower sampling rate that produced longer artifacts.

To remove artifacts, the signal was sampled at one instance between each RF transmission and linearly interpolated between these points. Recordings to BP+2 stimulation contained artifacts that lasted for 700 μs after each stimulation pulse onset. The BP+5 artifact lasted for 1000 μs. For consistent analysis, all artifacts were removed by interpolation at 1050 μs after stimulus pulse onset. The right-most columns of Fig. 4 in Hofmann and Wouters (2010) graphically illustrate the artifact removal technique.

After CI artifact rejection, a second-order butterworth high-pass filter attenuated frequencies below 2 Hz. To remove muscle potentials and other recording artifacts, the 10% of epochs with the highest peak to peak amplitude were rejected. To reduce variability and the influence of spontaneous neural activity, measurements from a subset of recording electrodes were averaged to give a single signal. Electrodes that were found to have significant responses as determined by a one-sample Hotelling T^2 test (Hofmann and Wouters, 2012) in more than 50% of recordings across subjects and stimulus electrodes were included in the average. Electrodes in the mirror image position of the head to those with 50% significant responses were also included. The selected electrodes were P9, P10, PO7, PO8, PO3, PO4, CP5, CP6, O1, O2, Oz, Iz and Pz. For each participant, any electrode that was touching the CI coil was removed from analysis. In all recordings, the phases of all selected electrodes were within ±30° of each other, ensuring there was no destructive interference. All results were calculated with Cz as the reference electrode.

A discrete Fourier transform was used to compute the complex frequency spectrum for each epoch. The response amplitude and phase were calculated as the mean of the response bin across epochs. Signal power was defined as the amplitude squared. The noise power was defined as the standard deviation of the response bin amplitude across epochs squared. The SNR was defined as the signal power in the frequency bin centred at the modulation frequency divided by the noise power in the same bin. A one sample Hotelling T^2 test, as used by Hofmann and Wouters (2012), was applied to determine whether EASSRs were present in the recorded signal.

2.5. Modulation detection

Modulation detection was assessed at 50% DR. Target stimuli were cathodic-first pulses amplitude modulated in cu at 20 Hz. The MDT studies mentioned in section 1 used a variety of different modulation rates including 10 Hz (Garadat et al., 2012; Zhou and Pfingst, 2014), 40 Hz (Pfingst et al., 2008) and 100 Hz (Fu, 2002). 20 Hz was used in this experiment as it was between the two EASSR rates and within the range used in previous CI studies. Modulation depth was defined as the peak-to-valley difference in cu. Measurements were taken at modulation depths of 1, 5, 10, 20 and 30 cu and 100% DR. If additional time was available, extra depths were included. If the peak-to-valley modulation depth was greater than 100% DR, then these depths were excluded. Such a wide range of modulation depths was tested to ensure that the MDT was reached for all participants.

As the loudness of fluctuating sounds in electrical hearing varies with modulation depth and absolute current level (McKay and Henshall, 2010), all modulated stimuli were loudness balanced to an unmodulated pulse train at 50% DR. For each modulation depth, a balancing procedure similar to that described in Section 2.4.1 was used. In the first trial the target was an unmodulated pulse train at 50% of the DR and the variable stimulus was modulated with a 20-Hz sinusoid with constant modulation depth. The listener was able to adjust the level of the entire modulated pulse train, while maintaining a constant modulation depth. Trials two through four followed the process previously described. Additionally, current level jitter of ±4 cu was applied to each stimulus during the MDT task to minimise any residual loudness cues after loudness balancing (Fraser and McKay, 2012).

A three-alternative forced-choice procedure was used to determine the modulation detection performance at each site and modulation depth. Three 500-ms stimuli were presented in each trial, separated by 500 ms of silence. Each stimulus was accompanied by a corresponding button changing color. Two intervals contained unmodulated stimuli, and one randomly selected interval contained the modulated stimulus. Participants were instructed to press the button corresponding to the interval that was different from the other two. 25 trials were performed at each depth and split across two measurement sessions to ensure continued concentration of the participant. Within each session, only a single stimulation site was tested. The order of presented stimulation sites and modulation depths was randomised. A training session of 10 trials at 100% modulation depth was completed for each site with feedback at the start of the experiment. For each stimulation site, psychometric functions were fitted to the percentage of correct responses as a function of modulation depth using the toolbox of Fründ et al. (2011). From the fitted curves the threshold was defined as the mid point between chance (33%) and completely correct (100%). The 66% correct point was extracted as the threshold, converted to ampere and reported in dB.

2.6. Statistical analysis

To determine the relation between behavioural MDTs and objective EASSRs, regression analysis was conducted on the twelve paired measurements. MDTs were compared to the SNR and amplitude of the EASSR. Analysis was conducted on the individual modulation frequencies and multiple regression analysis was used to assess the contribution from multiple modulation frequencies. In both the 4- and 40-Hz conditions, all 12 data points were included in the statistical analysis regardless of whether a significant response was detected. Statistical analysis was conducted using the R programming language.

3. Results

3.1. Threshold, comfort and stimulation levels

To elicit an equal loudness percept between unmodulated and modulated stimuli, $C_{m,b}$ was always greater than $C_{u,b}$. The difference varied between 2 cu and 19 cu (mean = 6.5 cu, sd = 4.4 cu). To balance anodic- with cathodic-first modulated stimuli, the additional current units required varied between −4.3 cu and 8.3 cu (mean = 0.5 cu, sd = 3.1 cu). Modulated stimuli at varying depths were loudness balanced to a constant pulse train. Fig. 1 illustrates the additional current units required to elicit an equal loudness percept for each of these conditions.
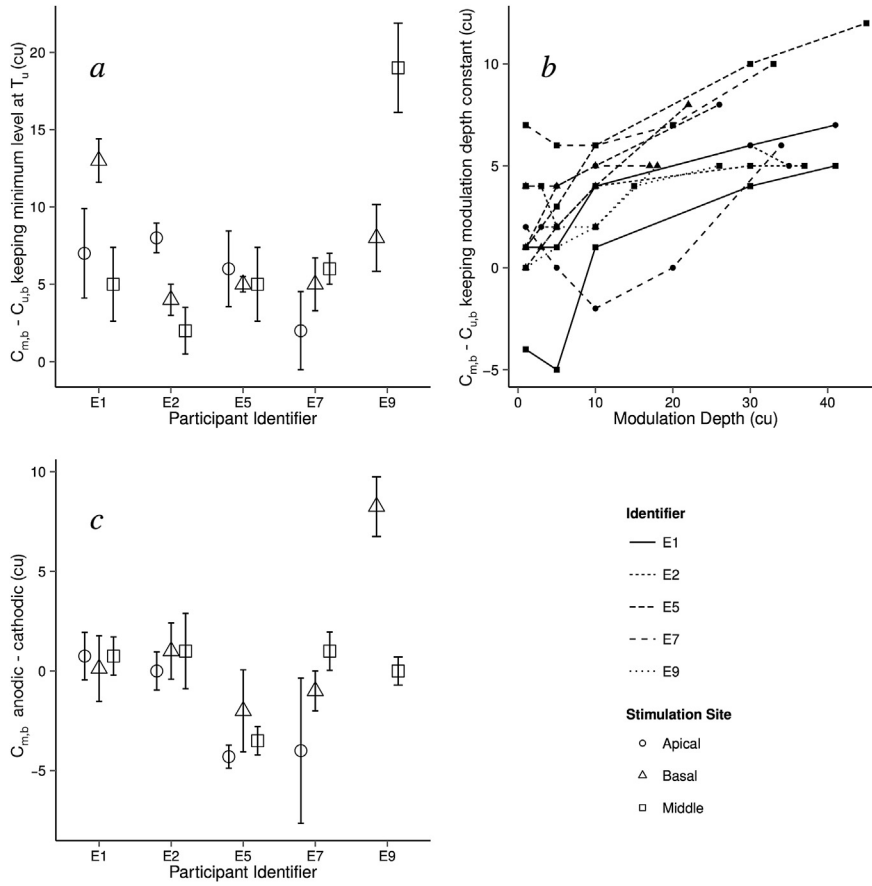
Fig. 1. Loudness balancing results. Difference in peak cu required to balance: a) unmodulated pulse trains to modulated stimuli with peak-valley modulation depth equal to 100% DR; b) modulated pulse trains at increasing modulation depth to an unmodulated pulse train; c) cathodic-first modulated pulse train to anodic-first modulated pulse train. Symbols indicate the stimulation site. Error bars extend ±1 standard deviation from the mean.

3.2. Electrically evoked auditory steady state responses

Fig. 2 shows the estimated EASSR amplitudes for all measurements in bipolar+2 mode when the signal was sampled at different times after stimulation. It can be seen that initially the signal was distorted and the amplitude artificially increased. This artifact peaked around 400 μs after stimulation, and decreased with further increase of time to 700 μs. Statistical analysis of the signal sampled before 700 μs incorrectly indicates every recording as containing a significant EASSR response. Similar analysis indicated that the bipolar+5 artifact lasted 1000 μs. All further results are for artifacts removed by interpolating the signal 1050 μs after stimulus onset. To confirm that no residual artifact was present, the false detection rate at neighbouring frequencies was checked to be 5%. Also the
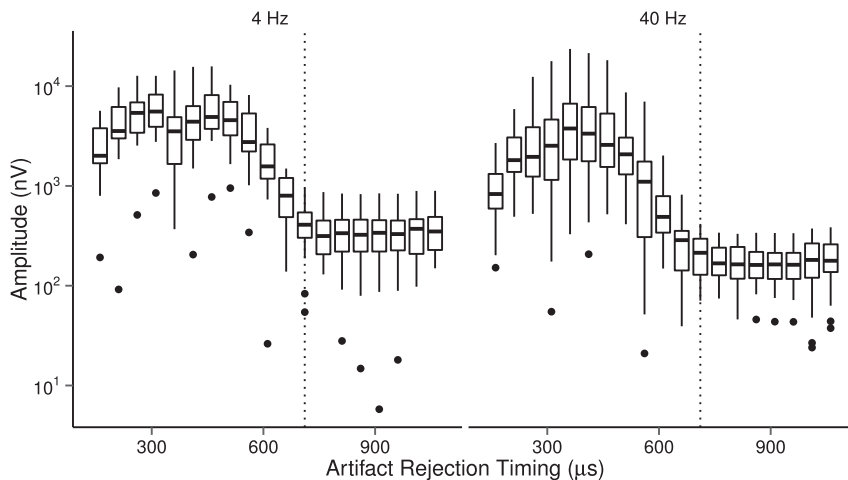

Fig. 2. Estimated amplitude of response for EASSRs sampled at different times after each RF transmission. Vertical dotted line indicates 700 μs after stimulation. Before this, the amplitude is dominated by the electrical artifact. After, the amplitude stabilises and reflects the neural response.

response phase was checked, to ensure it was not perfectly aligned with the RF transmission and electrical stimulation.

EASSRs were detected for all participants. For the 10-min measurements, the percentage of significant responses was 83% and 100% for the 4- and 40-Hz measurements, respectively. Measurements longer than 10 min per condition would reduce the noise level and possibly increase the detection rate at 4 Hz. On average, amplitudes were no greater for anodic-first than for cathodic-first stimuli. The difference was not significant for either modulation frequency: 4 Hz, t(11) = 1.05, p = 0.32; 40 Hz, t(11) = −1.11, p = 0.29. As there was no difference in responses for the two polarities, the two signals were merged as described in Section 2.4.3, and analysed as a single recording. All of the following results are for the merged signals.

A summary of the EASSR response and background characteristics is given in Table 2. This information is presented graphically in Fig. 3, which summarises the amplitudes, and Fig. 4, which summarises the phase and SNR for each modulation frequency. Background activity was lower for the higher modulation frequencies, and larger SNRs were found at 40 Hz than at 4 Hz.

3.3. Modulation detection

MDTs determined from the fitted psychometric curves (see Fig. 5 for an example) varied between −32 and −12 dB. The accuracy of the MDTs was limited by the number of balanced modulation depths, resulting in a large variation of the 95% confidence intervals.

3.4. Statistical analysis

Regression analysis was used to assess the relation between EASSRs and MDTs. 40-Hz EASSR SNRs were significantly related to MDTs, $F(1,10) = 16.19$, $R^2 = 0.62$, $p = 0.002$, as is illustrated in Fig. 6. Also, 40-Hz amplitudes were significantly related to the MDTs, $F(1,10) = 10.22$, $R^2 = 0.51$, $p = 0.009$. 40-Hz SNRs were not significantly more highly correlated with the MDTs than 40-Hz amplitudes, as determined by a Steiger test, $z = −0.90$, $p = 0.37$. 4-Hz EASSR SNRs were not significantly related to MDTs, $F(1,10) = 2.37$, $R^2 = 0.19$, $p = 0.15$. Also, 4-Hz amplitudes were not significantly related to the MDTs, $F(1,10) = 0.95$, $R^2 = 0.09$, $p = 0.35$.

Multiple regression was used to assess how well the 4-Hz and 40-Hz EASSRs predicted MDTs together. The SNRs explained 64% of the variance in MDTs, $F(2,9) = 7.92$, $R^2 = 0.64$, $p = 0.01$. The amplitudes explained 51% of the variance, $F(2,9) = 4.68$, $R^2 = 0.51$, $p = 0.04$. An ANOVA indicated that including the 4-Hz EASSRs provided no significant improvement in the prediction of MDTs over the 40-Hz data alone.

40-Hz EASSR SNRs from the anodic first and cathodic first stimuli were analysed independently and used to predict the MDTs. Both regression analyses were significant: anodic first, $F(1,10) = 19.04$, $R^2 = 0.66$, $p = 0.001$; cathodic first, $F(1,10) = 11.1$, $R^2 = 0.52$, $p = 0.007$. The Steiger test showed no statistical
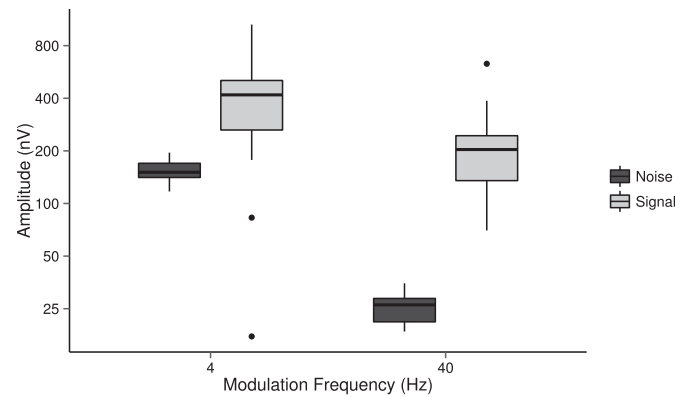
Fig. 3. Signal and background EASSR amplitudes for both anodic- and cathodic-first measurements. Boxplot edges correspond to the first and third quartiles, lines represent the median value, whiskers represent the largest value within 1.5 × the inter quartile range, and points outside this are marked as outliers.

difference between the anodic and cathodic EASSRs in predicting the MDTs, $z = −0.7$, $p = 0.45$.

4. Discussion

Both objective EASSRs and behavioural MDTs were measured for 12 cochlear implant electrode pairs in five participants. Regression analysis showed that the 40-Hz EASSR was significantly related with the MDT, confirming the hypothesis that EASSRs may be used to assess site-specific temporal sensitivity.

For CI users, temporal sensitivity is of considerable importance due to their reliance on low-frequency modulation cues in the speech envelope. The standard measure of temporal sensitivity is modulation detection, which is related to vowel, consonant and word recognition (Fu, 2002; Won et al., 2011). Furthermore, MAP adjustments based on MDTs, such as site removal or T level adjustment, have resulted in significant speech perception improvements (Garadat et al., 2012; Zhou and Pfingst, 2014).

Participants with a wide range of speech perception performance were recruited for this study in an attempt to characterise the entire breadth of temporal sensitivity outcomes. This approach resulted in a spread of MDTs of 20 dB which is illustrated in Fig. 6. But contrary to expectations (Fu, 2002), users with the best and worst speech performance did not correspond to those with the best and worst MDTs. Each modulated stimulus was loudness balanced to a reference signal to ensure that the task was not performed using loudness cues (Fraser and McKay, 2012). We found that to maintain equal loudness with an unmodulated pulse train, the peak of the modulated signal had to be increased with increasing modulation depth. Future experiments could interpolate the offset required to produce an equal loudness percept, which would allow the use of adaptive procedures and reduce measurement time.

Two primary factors hinder the use of MDTs as a clinical tool: (1) MDTs require an active participant for an extended period of time, which makes testing young children a challenge, and (2) MDTs require long measurement times. These constraints have motivated the desire for an objective measure to assess the temporal sensitivity of cochlear stimulation channels.

This study measured EASSRs at two modulation frequencies, both related to speech perception (Alaerts et al., 2009; Poelmans et al., 2012). 4 Hz is a prominent modulation frequency in the average envelope of speech (Assmann and Summerfield, 2004; Plomp, 1983; Edwards and Chang, 2013) and approximates the rate of syllables in running speech (Greenberg, 1999). Modulation
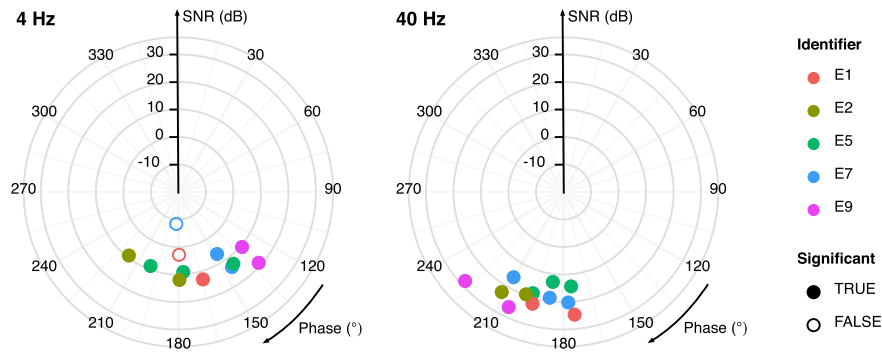
Table 2
EASSR response and background activity statistics. Median (Q2) and inter-quartile ranges (IQR) are listed for amplitude, SNR and phase, including both anodic and cathodic polarities. Phase statistics are calculated only for significant responses.

Freq (Hz)	Response						Background	
	Amp (nV)		SNR (dB)		Phase (°)		Amp (nV)	
	Q2	IQR	Q2	IQR	Q2	IQR	Q2	IQR
4	419	242	7.1	5.4	163	45	151	29
40	204	109	17.2	6.0	195	24	26	8

Fig. 4. Phase versus SNR for merged EASSRs. The angle represents the phase of the response and the distance from the centre represents the SNR of the response. Color indicates the participant identifier, filled circles indicate significant responses, and empty circles indicate that no significant response was detected. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
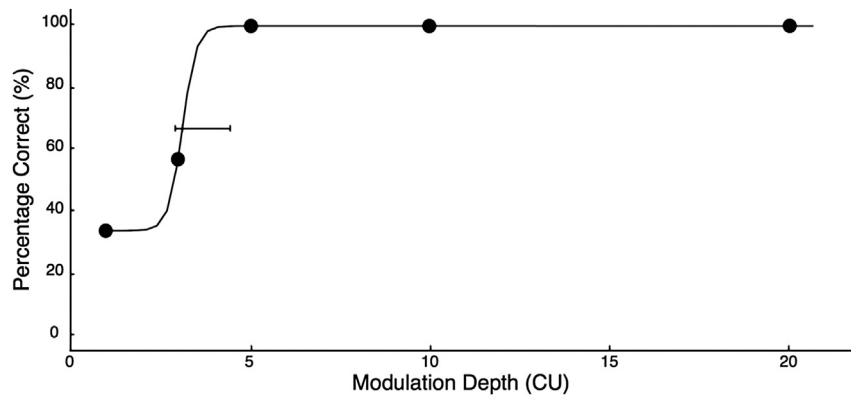

Fig. 5. Example fitted psychometric curve with 95% confidence interval for MDT determination. Asymmetric confidence intervals were a result of the experimental procedure. In this example an extra measurement was taken at depth 3 cu.

rates in the range of 40 Hz give the largest ASSR SNRs (Picton et al., 2003). Both modulation rates are within the range of perceptually relevant envelope modulation frequencies for CI users (Stone et al., 2008; Füllgrabe et al., 2009). Given the envelope extraction

Fig. 6. Objective and behavioural measurements. The abscissa represents the EASSR multiple regression model and the ordinate the MDT measurements. Error bars represent 95% confidence intervals.

strategies employed by most of the current generation CIs, and the extra reliance of CI users on temporal cues (Friesen et al., 2001), we hypothesised that the link between speech understanding and these low frequency ASSRs may be more pronounced in CI users than in normal hearing.

Consistent EASSR responses were detected for both modulation frequencies. Longer measurement times would decrease the noise level, and most probably increase the detection rate at 4 Hz. In acoustic hearing, the 40-Hz response is known to have a larger SNR than for other modulation frequencies (Picton, 2010). Similarly, we found that EASSR SNRs where higher at 40 than at 4 Hz.

A significant relation was found between objective and behavioural measures; as EASSRs increased the MDTs improved. This relation existed for the EASSR SNRs and amplitudes for the 40-Hz modulation data, and the combined 4- and 40-Hz data. The relation between EASSRs and MDTs was primarily driven by across-subject differences. Further investigation should be conducted of across-electrode variation within individual arrays. A procedure that produces smaller error in the MDT results would help to distinguish across-electrode differences.

The relation between EASSRs and MDTs was predominantly driven by the 40-Hz EASSR measurements. The higher SNRs at 40 that at 4 Hz may have provided a more accurate sampling of the underlying neural mechanisms and this may explain the stronger relationship with MDTs. The 40-Hz response is generated earlier in the auditory pathway than the cortically generated 4-Hz response. The former is generated closer to the auditory nerve and brainstem (Herdman et al., 2002) and it may better reflect the signal transmitted from the electrode-neuron interface. Further investigation
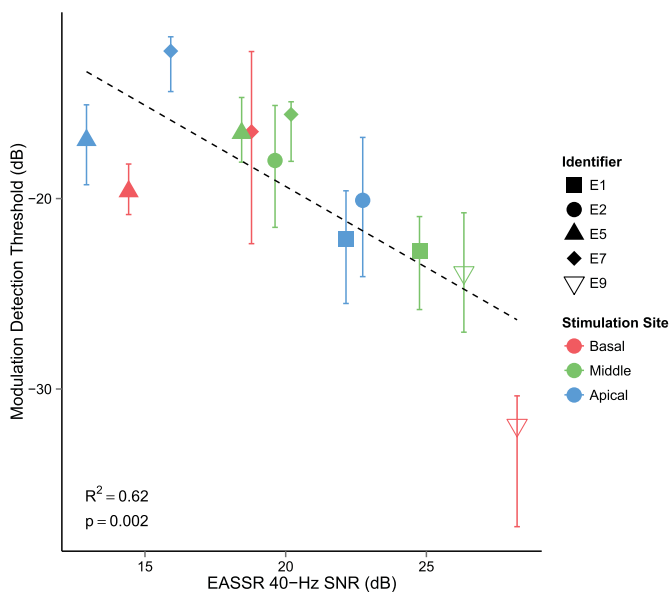
of the relation between EASSRs and temporal modulation sensitivity at higher modulation rates may determine if EASSRs generated earlier in the auditory pathway better predict MDTs.

The choice of electrical stimulation parameters was restricted by current EASSR artifact removal techniques. The implemented technique requires a balance between phase width, pulse rate and stimulation mode. Increased phase width and wider stimulation modes produce longer artifacts and faster pulse rates leave less time to sample the artifact free data. In this study we chose to use clinically relevant stimulation rates at the cost of using a non-standard stimulation mode. Bipolar stimulation mode was chosen as it may activate a more restricted region of the cochlea than monopolar stimulation (Snyder et al., 2008; Kwon & van den Honert, 2006) and act as a useful indicator of localised neural health (Chatterjee and Yu, 2010). Response amplitudes at 40 Hz were comparable to those for the bipolar measurements reported by Hofmann and Wouters (2010). However, there are no published reports of EASSR responses for frequencies below 40 Hz. When compared to acoustic responses, the 4-Hz amplitudes and SNRs were in the expected range (Picton et al., 2003).

Bipolar stimulation was also used in the modulation detection task to ensure that both measures were assessing the same cochlear space. The use of bipolar stimulation limits the immediate applicability of this study to commercial devices, as across-site MDT patterns differ depending on configuration mode (Pfingst, 2011). Improved artifact rejection techniques are required to measure EASSRs in monopolar mode. When these techniques become available, the relationship between EASSRs and MDTs should be re-evaluated with both the stimulation mode and rate matching those of commercial devices.

EASSR measurement for a single stimulation site took approximately five minutes using 40-Hz modulation and a single polarity. Further improvements could be made to reduce measurement time by optimising the stimulation parameters and post processing. ASSR techniques used to reduce measurement time could be adapted to EASSRs, including the multiple auditory steady-state response method (John et al., 1998) or advanced signal processing algorithms for noise removal and signal detection (De Cheveigné and Simon, 2008). Furthermore, other modulation frequencies may elicit larger and more reliable responses than at 4 and 40 Hz. The selection of modulation rates in this study was based on acoustic data of how responses vary with modulation rate (Picton et al., 2003; Alaerts et al., 2009). A study to determine which frequencies elicit the largest and most consistent responses in electrical hearing may reveal other more appropriate frequencies for modulation.

ASSRs are used as an early diagnosis tool for frequency specific hearing loss in hospitals. Their threshold detection reliability has been studied for both normal hearing and hearing impaired listeners (Luts and Wouters, 2005). The repeatability of ASSR amplitude measurements has also been studied (D'haenens et al., 2008). EASSRs are simply the electrical analogue, yet a systematic study of their reliability has not been conducted. Similarly, it should be investigated how long after implantation EASSRs become stable across the array, as more peripherally evoked potentials can take six to eight months to stabilise (Miller et al., 2008).

In conclusion it was found that EASSRs were significantly related to MDTs. As EASSRs increased, modulation detection improved. This indicates that EASSRs may be used as an objective measure of site-specific temporal sensitivity. Combined with measures of spectral sensitivity, this will allow for the practical characterisation of the electrode-neuron interface and provide an objective measure for the optimisation of site-specific stimulation parameters.

References

Alaerts, J., Luts, H., Hofmann, M., Wouters, J., 2009. Cortical auditory steady-state responses to low modulation rates. Int. J. Audiol. 48, 582—593.

Assmann, P., Summerfield, Q., 2004. The perception of speech under adverse conditions. In: Greenberg, S., Ainsworth, W.A., Fay, R.R. (Eds.), Speech Processing in the Auditory System. Springer, New York, pp. 231—308.

Bierer, J.A., 2007. Threshold and channel interaction in cochlear implant users: evaluation of the tripolar electrode configuration. J. Acoust. Soc. Am. 121, 1642—1653.

Bierer, J.A., Faulkner, K.F., 2010. Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves. Ear Hear. 31, 247—258.

Bierer, J.A., Faulkner, K.F., Tremblay, K.L., 2011. Identifying cochlear implant channels with poor electrode-neuron interface: electrically-evoked auditory brainstem responses measured with the partial tripolar configuration. Ear Hear. 32, 436—444.

Brown, C.J., Abbas, P.J., Gantz, B., 1990. Electrically evoked whole-nerve action potentials: data from human cochlear implant users. J. Acoust. Soc. Am. 88, 1385—1391.

Chatterjee, M., Yu, J., 2010. A relation between electrode discrimination and amplitude modulation detection by cochlear implant listeners. J. Acoust. Soc. Am. 127, 415—426.

De Cheveigné, A., Simon, J.Z., 2008. Denoising based on spatial filtering. J. Neurosci. Methods 171, 331—339.

D'haenens, W., Vinck, B.M., De Vel, E., Maes, L., Bockstael, A., Keppler, H., Philips, B., Swinnen, F., Dhooge, I., 2008. Auditory steady-state responses in normal hearing adults: a test-retest reliability study. Int. J. Audiol. 47, 489—498.

Dimitrijevic, A., John, M.S., Picton, T.W., 2004. Auditory steady-state responses and word recognition scores in normal-hearing and hearing-impaired adults. Ear Hear. 25, 68—84.

Dimitrijevic, A., John, M.S., van Roon, P., Picton, T.W., 2001. Human auditory steady-state responses to tones independently modulated in both frequency and amplitude. Ear Hear. 22, 100—111.

Edwards, E., Chang, E.F., 2013. Syllabic (2—5 Hz) and fluctuation (1—10 Hz) ranges in speech and auditory processing. Hear. Res. 305, 113—134.

Francart, T., Van Wieringen, A., Wouters, J., 2008. Apex 3: a multi-purpose test platform for auditory psychophysical experiments. J. Neurosci. Methods 172, 283—293.

Fraser, M., McKay, C.M., 2012. Temporal modulation transfer functions in cochlear implantees using a method that limits overall loudness cues. Hear. Res. 283, 59—69.

Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J. Acoust. Soc. Am. 110, 1150—1163.

Fründ, I., Haenel, N.V., Wichmann, F.A., 2011. Inference for psychometric functions in the presence of nonstationary behavior. J. Vis. 11, 1—19.

Fu, Q., 2002. Temporal processing and speech recognition in cochlear implant users. Neuroreport 13, 1635—1639.

Füllgrabe, C., Stone, M.A., Moore, B.C.J., 2009. Contribution of very low amplitude-modulation rates to intelligibility in a competing-speech task. J. Acoust. Soc. Am. 125, 1277—1280.

Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. 78, 2643—2647.

Garadat, S.N., Zwolan, T.A., Pfingst, B.E., 2012. Across-site patterns of modulation detection: relation to speech recognition. J. Acoust. Soc. Am. 131, 4030—4041.

Garadat, S.N., Zwolan, T.A., Pfingst, B.E., 2013. Using temporal modulation sensitivity to select stimulation sites for processor maps in cochlear implant listeners. Audiol. Neurotol. 18, 247—260.

Greenberg, S., 1999. Speaking in shorthand—a syllable-centric perspective for understanding pronunciation variation. Speech Commun. 29, 159—176.

Herdman, A.T., Lins, O., Van Roon, P., Stapells, D.R., Scherg, M., Picton, T.W., 2002. Intracerebral sources of human auditory steady-state responses. Brain Topogr. 15, 69—86.

Hofmann, M., Wouters, J., 2010. Electrically evoked auditory steady state responses in cochlear implant users. J. Assoc. Res. Otolaryngol. 11, 267—282.

Hofmann, M., Wouters, J., 2012. Improved electrically evoked auditory steady-state response thresholds in humans. J. Assoc. Res. Otolaryngol. 13, 573—589.

John, M.S., Lins, O.G., Boucher, B.L., Picton, T.W., 1998. Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. Int. J. Audiol. 37, 59—82.

Kwon, B.J., van den Honert, C., 2006. Effect of electrode configuration on psychophysical forward masking in cochlear implant listeners. J. Acoust. Soc. Am. 119, 2994—3002.

Lazard, D.S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P.H., Skarzynski, H., Schauwers, K., O'Leary, S., et al., 2012. Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. PloS One 7, e48739.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467—477.

Long, C.J., Holden, T.A., McClelland, G.H., Parkinson, W.S., Shelton, C., Kelsall, D.C., Smith, Z.M., 2014. Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. J. Assoc. Res. Otolaryngol. 15, 293—304.

Luts, H., Wouters, J., 2005. Comparison of MASTER and AUDERA for measurement of auditory steady-state responses. Int. J. Audiol. 44, 244—253.

McKay, C.M., Henshall, K.R., 2010. Amplitude modulation and loudness in cochlear implantees. J. Assoc. Res. Otolaryngol. 11, 101—111.

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—197.

Pfingst, B.E., 2011. Effects of electrode configuration on cochlear implant modulation detection thresholds. J. Acoust. Soc. Am. 129, 3908—3915.

Pfingst, B.E., Burkholder-Juhasz, R.A., Xu, L., Thompson, C.S., 2008. Across-site patterns of modulation detection in listeners with cochlear implants. J. Acoust. Soc. Am. 123, 1054—1062.

Pfingst, B.E., Xu, L., Thompson, C.S., 2004. Across-site threshold variation in cochlear implants: relation to speech recognition. Audiol. Neurotol. 9, 341—352.

Picton, T.W., 2010. Human Auditory Evoked Potentials. Plural, San Diego.

Picton, T.W., Dimitrijevic, A., van Roon, P., John, M.S., Reed, M., Finkelstein, H., 2001. Possible roles for the auditory steady-state responses in fitting hearing aids. In: Seewald, R., Gravel, J. (Eds.), A Sound Foundation through Early Amplification: Proceedings of the Second International Conference. Phonak AG, Basel, pp. 59—69.

Picton, T.W., John, M.S., Dimitrijevic, A., Purcell, D., 2003. Human auditory steady-state responses. Int. J. Audiol. 42, 177—219.

Plomp, R., 1983. The role of modulation in hearing. In: Klinke, R., Hartmann, R. (Eds.), Hearing Physiological Bases and Psychophysics. Springer, Berlin Heidelberg, pp. 270—276.

Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquière, P., Wouters, J., 2012. Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. Ear Hear. 33, 134—143.

Rance, G., Rickards, F.W., Cohen, L.T., De Vidi, S., Clark, G.M., 1995. The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials. Ear Hear. 16, 499—507.

Snyder, R.L., Middlebrooks, J.C., Bonham, B.H., 2008. Cochlear implant electrode configuration effects on activation threshold and tonotopic selectivity. Hear. Res. 235, 23—38.

Stone, M.A., Füllgrabe, C., Moore, B.C.J., 2008. Benefit of high-rate envelope cues in vocoder processing: effect of number of channels and spectral region. J. Acoust. Soc. Am. 124, 2272—2282.

Van Wieringen, A., Wouters, J., 2008. List and lint: sentences and numbers for quantifying speech understanding in severely impaired listeners for flanders and the netherlands. Int. J. Audiol. 47, 348—355.

Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M., Rubinstein, J.T., 2011. Acoustic temporal modulation detection and speech perception in cochlear implant listeners. J. Acoust. Soc. Am. 130, 376—388.

Xu, L., Thompson, C.S., Pfingst, B.E., 2005. Relative contributions of spectral and temporal cues for phoneme recognition. J. Acoust. Soc. Am. 117, 3255—3267.

Zhou, N., Pfingst, B.E., 2014. Effects of site-specific level adjustments on speech recognition with cochlear implants. Ear Hear. 35, 30—40.

Zwolan, T.A., Collins, L.M., Wakefield, G.H., 1997. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. J. Acoust. Soc. Am. 102, 3673—3685.