

Neural Encoding of Spectro-Temporal Cues at Slow and Near Speech-Rate in Cochlear Implant Users

Jaime A. Undurraga ^{*1}, Lindsey Van Yper¹, Manohar Bance², David McAlpine¹, and Deborah Vickers²

¹*Department of Linguistics, 16 University Avenue, Macquarie University, NSW 2109, Australia*

²*Department of Clinical Neurosciences, Cambridge Biomedical Campus, University of Cambridge, CB2 0QQ, UK*

The authors declare no competing financial interests.

*Corresponding author: Jaime A. Undurraga, jaime.undurraga@mq.edu.au

Abstract

The ability to process rapid modulations in the spectro-temporal structure of sounds is critical for speech comprehension. For users of cochlear implants (CIs), spectral cues in speech are conveyed by differential stimulation of electrode contacts along the cochlea, and temporal cues in terms of the amplitude of stimulating electrical pulses, which track the amplitude-modulated (AM'ed) envelope of speech sounds. Whilst survival of inner-ear neurons and spread of electrical current are known factors that limit the representation of speech information in CI listeners, limitations in the neural representation of dynamic spectro-temporal cues common to speech are also likely to play a role. We assessed the ability of CI listeners to process spectro-temporal cues varying at rates typically present in human speech. Employing an auditory change complex (ACC) paradigm, and a low (0.5 Hz) alternating rate between stimulating electrodes, or different AM frequencies, to evoke a transient cortical ACC, we demonstrate that CI listeners—like normal-hearing listeners—are sensitive to transitions in the spectral- and temporal-domain. However, CI listeners showed impaired cortical responses when either spectral or temporal cues were alternated at faster, speech-like (6-7 Hz), rates. Specifically, auditory change following responses—reliably obtained in normal-hearing listeners—were small or absent in CI users, indicating that cortical adaptation to alternating cues at speech-like rates is stronger under electrical stimulation. In CI listeners, temporal processing was also influenced by the polarity—behaviourally—and rate of presentation of electrical pulses—both neurally and behaviorally. Limitations in the ability to process dynamic spectro-temporal cues will likely impact speech comprehension in CI users.

1 Introduction

Since their inception, cochlear implants (CIs) have proven a transformative technology, restoring, or providing *de novo*, functional hearing and the ability to communicate using spoken language, in severe-profoundly deaf adults and children. CIs work by filtering acoustic signals into frequency (spectral) channels, and extracting the slowly varying amplitude envelope in each channel, which is then used to modulate a train of electrical pulses delivered by each electrode along the implanted array, directly stimulating the auditory nerve fibres (ANFs).

Despite their undoubted success, speech perception remains highly variable across the population of CI users (e.g. Holden et al., 2013). Potential sources of this variance are numerous, but likely include the quality of the electrode-neural interface (Pfungst et al., 2008; Bierer et al., 2010)—the degree of neural survival in the deafened cochlea, for example—or traumatic events arising from surgical insertion of the electrode array. A poor electrode-neural interface can result in broader neural activation than would otherwise be the case, as more current is required to excite ANFs. Further, stimulation parameters such as pulse rate, electrode configuration (monopolar *vs.* bipolar), and stimulus polarity can interact with the degree of neural health near to a stimulating electrode, affecting the overall fidelity of information encoded in ANFs (Macherey et al., 2008, 2017; Undurraga et al., 2010, 2012, 2013). Given the already reduced number of independent frequency channels in implanted, compared to normal-hearing (NH) ears, any additional distortions of the internal representation of the sound will contribute to reduced speech understanding (Eisen and Franck, 2005; Bierer, 2007).

CI listeners rely on spectral cues conveyed in terms of which electrode contacts along the inserted array are activated, and temporal cues conveyed in terms of the rate at which electrical pulses are modulated in amplitude, to understand speech. Since speech perception correlates with the number of electrodes accurately conveying temporal information (Croghan et al., 2017), and because different amplitude modulation (AM) frequencies convey distinct temporal features of speech (Rosen, 1992), the ability to discriminate *between* different AM frequencies represents a more ecologically valid approach to predicting speech perception outcomes in CI patients than simply detecting the *presence* of AM cues.

Here, we employed a measure of cortical function—the auditory change complex (ACC)—to assess neural representation of spectro-temporal cues in CI users. In contrast to the standard approach, in which a long, silent interval is presented between pairs of sounds, with the cortical potential at the transition between them—the ACC—reflecting neural/behavioural sensitivity to the physical/perceptual difference between the sounds (Fig. 1A), we employed an ‘alternating ACC’ paradigm in which stimuli are continuously and periodically alternated (i.e. without silent gaps), at a controlled rate such that every change triggers a cortical transient response (Fig. 1B) or a steady-state auditory change following response (AC-FR) (Fig. 1C).

[Figure 1 about here.]

We confirm that CI listeners are sensitive to slow alternating (0.5 Hz) spectral cues; like NH listeners, they can distinguish similarly alternating transitions in AM frequencies at the behavioural and neural levels, indicating no specific limitation to processing AM cues *per se*. As neural coding of AM cues may reflect interactions between stimulus polarity and the site of generation of action potentials in ANFs (Joshi et al., 2017), we also assessed the effect of stimulus polarity (anodic and cathodic) in CI listeners using a slow (126 pulses per second (pps)) pulse rate, and find that, behaviourally, anodic pulses are more distinguishable than cathodic ones, and neurally, cortical responses are weak regardless of polarity. Finally, employing speech-like (6-7 Hz) alternating spectro-temporal cues, we demonstrate a specific reduction in cortical responses in otherwise high-performing CI listeners, compared to NH listeners, indicative of greater neural adaptation at this higher alternating rate. Stronger cortical adaptation to speech-like rates of spectro-temporal modulations are likely the consequence of the greater neural synchrony of inner-ear neural responses to electrical stimulation, suggesting that current stimulation strategies might impose a fundamental limitation to speech performance outcomes in CI users.

2 Materials and Methods

Two experiments were conducted to measure neural responses to spectral or temporal cues. The first experiment was designed to measure neural responses to slow alternating cues (spectral or temporal) in an ongoing stimulus - a transient-like cortical ACC response. The alternating ACC paradigm was used at a low alternating rate (0.5 Hz), thus allowing us to minimize neural adaptation.

The second experiment was designed to trigger neural adaptation. Thus, the alternating paradigm was used at a higher rate (≈ 6 to 7 Hz), where neural adaptation is expected to be apparent. This alternating rate is interesting because it is close to the syllabic rate encountered in English and other languages, a rate to which the brain can entrain (4 to 7 Hz; Lewis, 2014; Peelle and Davis, 2012; Chandrasekaran et al., 2009). Hence, providing neural information at a rate critical for speech processing. In addition, this rate range is useful because it could be used to elicit a steady-state ACC response—the AC-FRs (Fig. 1C)—similar to those reported in NH studies using binaural stimuli (Dajani and Picton, 2006; Haywood et al., 2015; Undurraga

et al., 2016).

Note that, for simplicity, we use the term ACC, AC-FR, or auditory steady-state response (ASSR) indistinctly for acoustically- or electrically-evoked responses.

2.1 Participants

There were six NH listeners who were screened for NH, defined as pure-tone thresholds ≤ 20 dB HL, at octave frequencies from 250 to 8000 Hz; and five CI participants who were unilaterally implanted with the Cochlear Limited device and met the following inclusion criteria: (1) full intra-cochlear electrode insertion, (2) post implant activation experience of more than 9 months, and (3) post-lingual (after the acquisition of speech and language) onset of severe/ profound bilateral hearing loss.

Ethical approval was obtained from Macquarie University ethics committee (Ref. Number 5201700786) and by the Sydney Cochlear Implant Centre (SCIC) Cochlear Implant Program Research Governance Committee. All subjects provided informed consent before beginning the experiments and were paid an honorarium for their time.

2.2 Stimulus

With CI stimulation, stimuli consisted of the so-called quadra-phasic (QP) pulses (Macherey et al., 2017) where a positive (anodic) or negative (cathodic) pulse is emulated by concatenating two biphasic pulses (separated by $7 \mu\text{s}$) of opposite polarity. Therefore, to generate an “anodic” stimulus a cathodic-first biphasic pulse is concatenated to an anodic-first biphasic pulse so that the two central phases have the same anodic polarity—the quadra-phasic-anodic (QP-A) pulse (see Fig. 1E). A “cathodic” stimulus is generated in the same way by inverting the polarity—the quadra-phasic-cathodic (QP-C) pulse.

Spectral processing was assessed using the new alternating ACC paradigm we developed to detect neural responses to alternations between two fixed stimulating electrodes (Fig. 1B). This measure also served as a reference response since we had previously found reliable transient ACC responses using the standard ACC paradigm (Mathew et al., 2017, 2018). This was performed on three different loudness balanced (see Stimulus Presentation) electrode pairs, thus providing an objective assessment of electrode discrimination whilst minimizing the contribution

of loudness differences in the response.

Stimuli consisted of QP-C monopolar pulse trains at a rate of 827 pps, 43 μ s pulse width and 7 μ s inter-phase gap (IPG). For 4 out of 5 participants three apical electrode pairs were tested: 14-15, 15-16, and 16-17. For one participant (S4) the electrode pair 21-22 was used instead of 16-17 (see [Table 1](#) for details). Middle-apical electrodes were preferred as they convey most of the speech information by the CI, and also for their importance in speech perception (Busby and Clark, 2000; Geier and Norton, 1992; Henry et al., 2000).

Transient ACC responses were evoked using an alternating rate of 0.5 Hz. Each recording lasted 10 minutes and data were epoched in 300 blocks of 2 seconds. Adapted responses, measured by the AC-FR, were elicited using the alternating paradigm at a rate of 6.1 Hz. Each electrode pair was tested for about 5.43 minutes and data were epoched in 84 blocks of 3.92 seconds. Note that longer epochs for AC-FR analyses improve the spectral resolution in the frequency-domain (0.25 Hz resolution).

Temporal processing employed the alternating ACC paradigm to detect neural responses elicited by alternations between two fixed AM frequencies ([Fig. 1D](#)). We measured responses from both NH and CI listeners at two modulation depths, which was the same for the two AM frequencies—a paradigm that reduces the potentially confounding contribution of loudness cues to envelope encoding in CI listeners (McKay and Henshall, 2010); loudness perception is almost constant as a function of AM frequencies (Gransier et al., 2016; Kreft et al., 2010). An advantageous feature of this paradigm is that in addition to evoking the ACC response, the ASSR to the AMs can be obtained simultaneously. This is an indication of the phase locking ability in the auditory system (Picton et al., 2003), thus providing additional information about AM processing.

For the NH listeners a 500 Hz sinusoidal carrier was used and amplitude-modulated (AM'ed). AM frequencies alternated between of 20.0 and 35.5 Hz.

Two modulation depths were used in different conditions, but did not change within a condition. These were fully modulated (100%), and halfway modulated (50%, [Fig. 1D](#)). The change of AM frequency was always performed at the minima of the AM cycle to avoid other cues (e.g. clicks) occurring that would elicit an exogenous response.

For the CI listeners, stimuli AMs were presented to electrode 16. This middle-apical electrode

was selected because animal data indicate that apical neurons may encode temporal cues more strongly than basal neurons (Middlebrooks and Snyder, 2010).

QP pulses were AM'ed between 20 and 35 Hz, except for S5, where AMs alternated between 20 and 28 Hz (subject S5 could also discriminate this condition). Note that, in all cases, the magnitude of the AM frequency change was above discrimination limens reported in previous studies ($\approx 12\%$ of the base frequency for CI and $\approx 8\%$ for NH listeners; e.g. Chen and Zeng (2004) and Kreft et al. (2013)).

As with NH listeners, two AM depths were used at either 100 or 50% of the subject's dynamic range (in Current Units), defined as the difference between the comfortable and the threshold level (see Table 1 for details).

Two stimulation pulse rates 801 pps (4 subjects) and 126 pps (5 subjects) were used to encode the AM envelopes. We chose a low stimulation rate of 126 pps for two reasons. First, by using a low pulse rate, the effect of stimulus polarity on ANFs is maximized as the interactions between adjacent pulses decrease with increasing inter-stimulus interval (ISI) (Macherey et al., 2008; Macherey and Carlyon, 2012; Macherey et al., 2017; Undurraga et al., 2010, 2012, 2013). Second, this stimulation rate is suitable for reliably removing the electrical artefact from the ASSR via interpolation (Hofmann and Wouters, 2010; Luke et al., 2015; Deprez et al., 2017), hence, allowing us to estimate the amount of phase-locking to each AM frequency by means of the ASSR which is elicited alongside the ACC response.

The effect of polarity was investigated at 126 pps using positive (QP-A) and negative (QP-C) pulses, and all stimuli were checked in advance to ensure that they were discriminable at 100% modulation depth for the QP-A pulse shape.

At 801 pps, we only used QP-C pulses mainly because the short ISI between QPs pulses results in a "pseudo alternating" polarity pulse train. This can be observed in Fig. 1E, where the reduction of the ISI brings the last phase of the first QP pulse closer to the first phase of the second QP pulse. For both NH and CI listeners, transient ACC responses were evoked by alternating the AM at a rate of 0.5 Hz. Each recording lasted 10 minutes and data were (300 epochs of 2 seconds).

Neural adaptation to rapid AM frequency alternations was measured for both NH and CI listeners. For NH listeners, AC-FRs were measured to AM frequencies alternating between 19.2

and 32.0 Hz at a rate of at 6.4 Hz (5.02 mins, 69 epochs of 4.38 seconds, 0.23 Hz resolution). This slight difference in AM frequencies (compared to AM frequencies used in the low alternating rate) was necessary to ensure that an integer number of modulations fitted within each epoch. In order to elicit AC-FRs to AM frequency alternations in CI users, we used an alternating rate of 6.9 Hz (4.81 mins, 72 epochs of 4.01 seconds, 0.25 Hz resolution). It should be noted that the alternating rate was adjusted to ensure a fix integer number of pulses and AM cycles within an epoch which also depended on the stimulation pulse rate. This resulted in a slightly higher alternating rate than for the NH listeners (0.5 Hz difference). However, in all cases, the alternating rate was confined within the optimal range in which the AC-FR has been obtained using binaural paradigms (\approx 6 to 8 Hz; e.g. Dajani and Picton, 2006; Haywood et al., 2015; McAlpine et al., 2016; Undurraga et al., 2016).

A summary including pulse rate, AM frequencies and depths, pulse shape, electrodes tested, and different alternating rates for both spectral and temporal processing is provided in [Table 2](#).

2.3 Stimulus Presentation

For CI listeners, all stimuli were presented at a loudness balanced comfortable level estimated using a loudness scale chart. Comfortable levels were obtained for each electrode, pulse rate, and polarity by gradually increasing the stimulation level until participants reported a loudness of 6 on a 10-point loudness scale chart. This procedure was repeated twice and the final balanced comfortable level corresponded to the average of these two measures. For the AM stimuli, comfortable levels were determined only for the loudest condition, i.e. 50% AM depth. However, all participants ranked the fully AM'ed stimuli as 6 in the loudness scale chart. Thresholds for the AM stimuli were obtained using unmodulated pulses in a similar way but set as the level corresponding to a loudness of 1 (“barely audible”) on the loudness scale chart.

For CI listeners, the stimuli were directly presented to the CI using the Nucleus Implant Controller (NIC) version 3 via the RF-Generator-XS (Cochlear Limited, Sydney, Australia) using a custom interface developed in Python. For NH listeners, acoustic stimuli were presented via earphones on the right ear (Etymotic ER2) at 72 dB SPL. All acoustic stimuli were generated in a custom interface “[aep_gui](#)” developed in Matlab and delivered via RME Fireface UC at a sampling rate of 48 kHz and calibrated using a 2 cc GRASS artificial ear.

2.4 EEG recordings

Electroencephography (EEG) recordings were made using a 64-channel high resolution BioSemi ActiveTwo (Amsterdam, The Netherlands) at a sampling rate of 16384 Hz and a resolution of 24 bits/sample (31 nV LSB). Channels were arranged according to the international 10–20 system. Two additional channels were placed on the left and right mastoid and eye movements were recorded with right infra-orbital and right lateral canthus channels. The voltage offset was always below ± 40 mV and usually within ± 20 mV.

For NH listeners, recordings were referenced to Cz. With CI users, scalp channels around the CI receiver package were not connected (typically 1-3 scalp electrodes) and responses were referenced to the contralateral mastoid.

2.5 EEG analysis

EEG analysis was as follows: Firstly, data was referenced to remove line noise. Afterwards, unconnected electrodes (those near the CI) or noisy EEG electrodes due to poor contact (usually electrodes with voltage offsets greater than 40 mV or strong DC components) were automatically detected and removed from the analysis.

For the low pulse rate recordings (126 pps), CI artefacts were removed by interpolation between points with minimum artefact contamination (a detailed description is provided in the Appendix). For transient responses, time-domain data were down sampled to 1024 Hz and band-pass filtered between 2 and 30 Hz (zero-phase, third-order finite impulse response (FIR) Kaiser filter). Eye blink artefacts were removed using a template matching suppression method (Valderrama et al., 2018).

Next, data were epoched and residual CI artefacts were identified and removed using denoising source separation (DSS) (details provided in the Appendix; de Cheveigné and Simon (2008) and Mathew et al. (2017)). Per-channel averages were obtained using a weighted average (full details provided in Mathew et al. (2017)). The objective detection of an ACC response and signal-to-noise ratio (SNR) estimations utilised the Hotelling's t-squared test (Hotelling-T², Golding et al. (2009) and Mathew et al. (2017)). The significance of the response was evaluated using a time window spanning between 50 and 400 ms relative to stimulus change. The P1-N1-P2 peaks and latencies were automatically detected within this time window by obtaining the minimum

(N1) and maxima before and after N1. For AC-FRs, the analysis pipeline was similar with the exception that data was not down-sampled and only a high pass filter (at 2 Hz) was applied. The objective detection of a frequency-domain response was also determined using Hotelling-T2 test. In this case, confidence intervals are estimated from the two-dimensional distribution of the frequency bin of interest (Picton et al., 1987, 2003). To reduce the dimensionality of the data existing across all EEG electrodes, we estimated the global field power (GFP), which corresponds to the spatial standard deviation of the EEG data across all recording electrodes (Skrandies, 1990). For the transient ACC response, the GFP estimation was modified. The modification consisted of computing the standard deviation of the N1-P2 amplitudes across channels rather than taking the peak amplitude of the waveform resulting by computing the standard deviation across all channels, thus taking into account the contribution of both peaks, and not just the main peak as with the standard GFP.

The latency of the steady-state responses was estimated following a series of adjustments previously described (Rodriguez et al., 1986; Herdman et al., 2002; John and Picton, 2000; Undurraga et al., 2016).

First, the phase (ϕ) of the frequency of interest (f_i) was adjusted by π rad when data were referenced to Cz and by 0 rad when referenced to the contralateral mastoid. This phase adjustment accounts for polarity differences arising the location of the reference electrode. Next, the phase of responses generated by the AM were adjusted by $\pi/2$ rad in order to compensate for the frequency transformation, which was computed in terms of a cosine phase. AC-FRs generated by rapid alternating cues did not include this correction as it was assumed that the response followed each transition, i.e. the transition had zero phase. Once the phase was adjusted (ϕ_a), the phase delay (ϕ_d) was computed as:

$$\phi_d = 2\pi - \phi_a \quad (1)$$

and the latency (in seconds) was estimated by:

$$L = \frac{1}{2\pi \cdot f_i} (\phi_d + 2\pi \cdot n + 2\pi \cdot m) \quad (2)$$

where n is an integer number set to correct for circularity ambiguity (phase unwrapping)

associated to phase and m is an integer number set to account for the number stimulus cycles occurring before a response is evoked. Here, m was set to zero when studying AC-FRs whilst it was set to 1 when analysing latency of the ASSRs generated by the AMs.

All signal processing was carried out off-line using a custom analysis module “[pyeeg-python](#)” developed in Python 3.

2.6 Behavioural discrimination

In all experiments, the ability of CI participants to behaviourally discriminate either between AM frequencies or between electrodes was assessed with a 3-interval 2-alternative forced choice task. Each interval lasted 400 ms and there was a 500 ms silent gap between intervals. The first interval always contained the reference stimulus, and the second and third contained either the reference or the target, and the participant indicated whether the 2nd or 3rd presentation was different to the other two. A total of 20 trials were presented for each pair of contrasts and based on binomial significance levels ($p < 0.01$) the discrimination threshold was set at 80%. The stimuli used in the behavioural discrimination were identical to those used to measure the brain response.

[Table 1 about here.]

[Table 2 about here.]

2.7 Statistical analyses

Analysis of variance (ANOVA) were carried out by means of linear mixed-effects (LME) models for which the factor subject was set as random. LME models with more than one factor were fitted and reduced using a backward stepwise reduction method. The degrees of freedoms were estimated by means of Satterthwaite approximation (Kuznetsova et al., 2017). If outliers were detected these were removed and the LME model was refitted. Conditional, random, and marginal residuals of the LME model were checked by visual inspection. For all LME models, the variance explained by the fixed factors, marginal R^2 (R_m^2), and by fixed and random factors, conditional R^2 (R_c^2), are reported as a tandem ($R_{m/c}^2$; Nakagawa and Schielzeth (2013) and Harrison et al. (2018)). Post-hoc comparisons were carried using trimmed means when heteroscedasticity was observed (Field and Wilcox, 2017; Mair and Wilcox, 2020).

All statistical analyses were performed using the R software package (R Development Core Team, 2020) and the significance level was set to $\alpha = 0.05$.

3 Results

3.1 Behavioural discrimination of temporal cues is affected by pulse polarity in CI listeners

We first examined the behavioural ability of CI listeners to discriminate between stimulation of different electrodes (spectral discrimination) and different AM frequencies (temporal discrimination) using high rate cathodic (QP-C) pulses, and low rate cathodic (QP-C) and anodic (QP-A) pulses. Overall, CI listeners performed well in both listening tasks (Fig. 2), and only one participant (S4) was unable to discriminate between stimulation of the most apical electrode pair (Fig. 2A). Similarly, in the AM discrimination task, although discrimination using the high pulse rate (801 pps) was at ceiling for all participants (Fig. 2B), sensitivity to AM cues using the low pulse rate (126 pps) was more variable (Fig. 2C). To determine which factors might explain the variability in performance across listeners at the lower (126 pps) pulse rate, a generalized LME model (using a logit family) was fitted to the data, with the significance of the response as the dependent variable. The model including factors *modulation depth* and *polarity* was significantly better than the null hypothesis model [$\chi^2(2) = 6.12$, $p = 0.046$, $R_{m/c}^2 = 0.427/0.654$]; in only one case (S5, Fig. 2C) the AM frequency was not discriminable with QP-A pulses, whilst this occurred for four cases using QP-C pulses.

As well as differences between listeners, we also observed differences within CI listeners in their ability to discriminate AM frequencies. Subjects S1 and S3 could discriminate between the two AM frequencies for both polarities and modulation depths, whilst subject S5 could only discriminate AM frequencies at 100% modulation depth for both polarities. Subject S2 could discriminate only QP-A pulses for both modulation depths, and S4 could discriminate all conditions except for QP-C pulses at 50% modulation depth.

Overall, our data demonstrate that CI listeners were able to discriminate differences in spectral cues elicited by stimulation of different electrode pairs, and that discrimination of AM cues varies as a function of the electrical pulse rate at which the cues are conveyed, as well the depth

of modulation and the pulse polarity.

[Figure 2 about here.]

3.2 CI users show robust neural discrimination between electrodes during ongoing stimulation

We next assessed cortical sensitivity to spectral cues in CI listeners using our alternating ACC paradigm, alternating periodically, at a slow rate (0.5 Hz), the stimulated electrode of an adjacent pair of electrodes (Fig. 1B). As with the standard paradigm (Mathew et al., 2017, 2018; He et al., 2014; Brown et al., 2008) neural responses to the electrode discrimination task evoked large ACC responses. Typical neural responses and topographic maps from the three electrode pairs confirm that electrical artefacts were successfully attenuated from the recordings (see Appendix for details), evidenced by the absence of a large area of activation on the side nearest the implanted (left) ear (Fig. 3, for CI participant S5). Overall, transient ACC responses were obtained from all participants and for all electrode pairs (Fig. 4) using the alternating ACC paradigm. To facilitate description of the EEG data, we selected the channel with the largest SNR for each participant. Across participants, F-values estimated via Hotelling-T2 test ranged between 4.38 and 144.79 (with a critical F for a significance level of $p < 0.05$ of 1.97). A LME model with factor *electrode pair* indicated that neither the GFP [$F(2, 8) = 3.87$, $p = 0.07$, $R^2_{m/c} = 0.033/0.941$] nor N1 latency [$F(2, 8) = 0.21$, $p = 0.81$, $R^2_{m/c} = 0.006/0.787$] were explained by this factor. Across subjects and electrodes, N1-P2 amplitudes ranged between 1.89 and 10.39 μV (s.d. = 2.80 μV), whilst N1 latencies spanned 93.0 to 207.5 ms (mean = 122.1 ms, s.d. = 33.3 ms). There was good correspondence between behavioural and ACC responses, with 14 out of 15 agreements between successful behavioural discrimination and objective detection of the response. Note that participant S4, who showed poor electrode discrimination for one electrode pair in the behavioural task (Fig. 2A), nevertheless showed strong ACC responses for all three electrodes comparisons. This is consistent with our previous study where we found that, for some participants, the ACC response preceded the ability to discriminate electrodes in a behavioural task (Mathew et al., 2018), and suggest that electrode discrimination may be limited at a more central (maladapted) stage in the brain.

The data confirm that periodic electrode alternations at a low rate in an ongoing and continuous stimulus evoke reliable ACC responses (Mathew et al., 2017, 2018).

[Figure 3 about here.]

[Figure 4 about here.]

3.3 Neural temporal discrimination differs between NH and CI listeners

Sensitivity to AM, and specifically to changes in AM frequency, is a critical factor in speech understanding. To assess the dynamics of AM sensitivity, we measured the cortical response to transitions in AM frequency in both NH and CI listeners, employing the transient ACC evoked by a low, alternating rate (0.5 Hz) in AM frequency (alternating between ≈ 20 and 35 Hz; Table 1).

In NH listeners, reliable cortical activation to transitions in AM frequency—assessed in terms of the GFP of transient ACC responses—were evoked at 100% AM depth (Fig. 5A), and were significantly smaller [$F(1, 5) = 24.05$, $p = 0.004$, $R_{m/c}^2 = 0.452/0.793$] at 50% AM depth (Fig. 5B & E). No significant effect was observed in terms of the latency of the transient ACC as a function of modulation depth (Fig. 5F). Across NH listeners, Hotelling-T2 F-values ranged between 3.8 and 8.6 for 50% modulation depth and 6.1 and 37.3 for 100% modulation depth. In addition, significant ASSRs were simultaneously obtained from all NH listeners at both 50% and 100% AM depths (Fig. 5C & D), and, for these, we observed a significant interaction between the AM frequency and modulation depth [$F(1, 15) = 4.79$, $p = 0.045$, $R_{m/c}^2 = 0.179/0.610$]. Overall, 35 Hz ASSRs were significantly larger at 100%, than at 50% AM depth [$t(5) = 6.34$, $p = 0.001$, $d' = 0.38$], whilst 20 Hz ASSRs were not significantly different as a function of AM depth. These results suggest that, for NH listeners, ACC responses are highly sensitive to AM depth, and that ASSRs, whilst affected by the AM depth, are able to encode AM frequencies at both depths.

[Figure 5 about here.]

Like NH listeners, CI listeners also showed significant alternating ACC responses to dynamic AM cues, but this sensitivity was highly dependent on pulse rate.

For the high pulse rate of 801 pps (presented as QP-C pulses) ACC responses were robust; in agreement with the behavioural data (Fig. 2B), transient ACC responses in 4 CI subjects (S5 was unable to take part in this experimental session) were clearly evoked to both AM modulation depths at this higher pulse rate (Figures 6A & B). This indicates that neurally, as well as behaviourally, the two AM rates are reliably encoded in CI listeners. In contrast to NH listeners, however, the magnitude of the transient ACC response was not significantly affected by modulation depth; a LME model indicated that the factor *modulation depth* was not significant for the GFP [$F(1, 1.93) = 0.19, p = 0.699, R^2_{m/c} = 0.006/0.851$] or for the latency [$F(1, 2.08) = 0.76, p = 0.471, R^2 = 0.014/0.903$] (Figures 6M & N). Hotelling-T2 F-values ranged between 10.3 and 38.6, comparable to the range observed in NH listeners. We compared the GFP and the latency of ACC responses between NH and CI using a LME model with factors *modulation depth* and *group* (NH or CI) as independent variables. The analysis indicated a significant interaction between *modulation depth* and *group* [$F(1, 7.20) = 8.94, p = 0.009, R^2_{m/c} = 0.449/0.813$], confirming that ACC responses for the two modulation depths differed for NH and CI listeners. Latencies, however, were not significantly affected by modulation depth. The mean N1 latencies were 182 ms (s.d. = 62 ms) and 164 ms (s.d. = 17 ms) for NH and CI listeners, respectively.

Next, we compared ACC responses generated by alternating between different electrodes (spectral alternations) and between different AM frequencies (temporal alternations) at 801 pps pulse rate in CI listeners. A LME model with the factor *type* (electrode and AM) indicated that the GFP [$F(1, 14.01) = 54.52, p < 0.001, R^2_{m/c} = 0.115/0.965$] and latency [$F(1, 15.11) = 34.35, p < 0.001, R^2_{m/c} = 0.428/0.768$] differed significantly, with ACC responses evoked by alternating electrodes (stimulated at 827 pps) around 50% larger, and latencies 42ms shorter, than those evoked by AM changes (stimulated at 801 pps).

[Figure 6 about here.]

In contrast to the 801 pps pulse rate, the 126 pps pulse rate showed a relatively poor correspondence between behaviourally sensitivity (Fig. 2C) and transient ACC responses, with only 2 out of 10 agreements when anodic (QP-A) pulses were employed. This was improved only marginally (4 out of 10 agreements) when cathodic (QP-C) pulses were employed. Overall, behavioural discrimination was more sensitive than ACC to changes in AM frequency; transient

ACC responses for both QP-A (Fig. 6E & F) and QP-C (Fig. 6I & J) pulse trains, respectively, across all subjects, revealed to be relatively weak and variable. Indeed, for QP-A pulses, average traces were dominated by the responses of S3 at 100% (Fig. 6E), and S3 and S5 at 50% (Fig. 6F) modulation depth, whilst for QP-C average traces were dominated by responses from S2 and S3 at 100% (Fig. 6I), and S1, S2, S3 at 50% (Fig. 6J) modulation depth. Similarly, ASSRs (embedded on the transient waveform) to the two AM frequencies were clearly observed in the frequency domain for S2 and S3, whose responses dominate the average waveform (Fig. 6G, H, K, & L). Across all participants and polarities, Hotelling-T2 F-values ranged between 2.6 and 18.1 for 50% modulation depth and 2.0 and 18.4 for 100% modulation depth (with a critical F of 1.97 for a significance level of $p < 0.05$). Although this analysis suggests that significant transient ACC responses were obtained from all participants, visual inspections revealed that some significant responses were not evoked by transitions in the AM frequency but, rather, by the 20 Hz ASSR component present in the response (Appendix ??). A LME model confirmed that neither modulation depth nor polarity had a significant effect on the GFP or the latency of response (Fig. 6O-R).

The data suggest that low pulse rates do not elicit a strong ACC and, consequently, are relatively insensitive to polarity effects. A plausible explanation for this is that the perceptual change resulting from transitions in the AM frequency is relatively weak due to the low carrier-rate used to convey the AMs; the 126 pps pulse train is only about 4 times higher than the highest AM rate we employed. This is reflected in the relative increase in variability of behavioural responses (cf. Figures 2B & C). Thus, whilst the stimuli are behaviourally discriminable, neural synchronization and/or activation may be reduced, leading to poor transient ACC responses. Overall, the data confirm that, as in NH listeners, the ACC is sensitive to alternating AM frequencies in CI users, but that this (and behavioural) sensitivity, depends upon the specific choice of electrical pulse rates applied in the CI processor; poor cortical sensitivity to the low (126 pps) pulse rate was not dependent on pulse polarity. Further, despite high (801 pps) pulse rates evoking large cortical ACC responses in CI listeners, these responses were insensitive to changes in AM depth, and cortical responses to alternating AM frequencies (temporal cues) were significantly smaller, and of longer latency, than responses to alternating electrodes (spectral cues).

3.4 CI listeners show hyper-adapted cortical responses to speech-like alternating cues

While CI listeners show robust ACC responses to slow spectral and temporal transitions, it is unclear whether these cues remain robust at more realistic (i.e. speech-like) alternating rates. Here, we re-assessed neural encoding of spectral and temporal cues by presenting them at a higher alternating rate (6 to 7 Hz) to evoke AC-FRs (Fig. 1C).

We first recorded AC-FRs to alternations in AM frequency in NH listeners at a rate of 6.4 Hz (Fig. 7A-D). At this alternating rate, clear AC-FRs were observed (see peaks in frequency-domain responses in Fig. 7C-D), although ASSRs to the AM frequencies were much smaller, likely because the short alternating period (156 ms), which is insufficient to generate substantial ASSRs in NH listeners (Ross et al., 2002). Significant responses were obtained from all participants at 100% AM depth, with Hotelling-T2 F-values ranging between 3.25 and 9.0 (critical F value 3.14 for $p < 0.05$). A LME model indicated that both *frequency* [$F(2, 6.01) = 7.99, p = 0.020, R_{m/c}^2 = 0.508/0.723$] and *modulation depth* [$F(1, 22.00) = 7.01, p = 0.015, R_{m/c}^2 = 0.508/0.723$] had a significant effect on the GFP (Fig. 7E). The mean latency of the steady-state responses (Eq. 2; Fig. 7F) generated by the alternating rate (6.4 Hz) between the two AM frequencies, and the AM frequencies in itself (19.2 and 32 Hz) were 90.0, 65.8, and 44.8 ms, respectively. Latencies of the responses to 6.4 Hz (where responses were reliably obtained) were not affected by the modulation depth. These data are consistent with those from the low alternating rate and demonstrate that AC-FRs to AM frequency alternations could be obtained from all NH listeners at rapid alternating rates, in a similar way to those evoked with binaural stimuli (Dajani and Picton, 2006; Haywood et al., 2015; McAlpine et al., 2016; Undurraga et al., 2016).

[Figure 7 about here.]

In contrast to NH listeners, steady-state cortical responses in CI listeners were considerably less robust to speech-like alternating rates. We assessed whether AC-FRs in CI listeners could phase-lock to the rapidly alternating AM frequencies (6.9 Hz), for pulse rates of 801 or 126 pps. However, in comparison to the robust ACC generated by the low (0.5 Hz) alternating rate using the 801 pps pulse rate, this same pulse rate did not generate a significant AC-FR; ASSRs were evident at the AM frequencies, but not AC-FRs at the 6.9 Hz alternating rate (Fig. 8A-D). Note

that the magnitude of the AC-FRs did not change when data were reanalyzed without applying spatial filtering for artifact rejection.

Unsurprisingly, at a pulse rate of 126 pps, AC-FRs were largely absent for either anodic or cathodic pulse shapes (Fig. 8E-L), consistent with the transient ACC responses obtained under the same conditions. In contrast, relatively strong ASSRs were observed at the two AM frequencies (Fig. 8, vertical lines in Fig. 8G, H, K, & L) as well as comodulations around those rates, as demonstrated by the spectral peaks near the main AM frequencies. Comodulations were caused by the beating of each AM frequency within the analysis window used to compute the frequency-domain response (4.01 seconds). Individually, each AM frequency is effectively presented at a rate of 3.45 Hz (half the alternating rate), resulting in a temporal beating which is equivalent to a second order modulation observed around each AM frequency ± 3.45 Hz.

The data confirm that, for CI listeners, ASSRs are observed, even when the duration of each AM frequency block (≈ 145 ms) is close to the minimum time required to evoke ASSRs in NH listeners (Ross et al., 2002). However, consistent with an excess of cortical neural adaptation in CI listeners, AC-FRs were smaller than in NH listeners at this alternating rate (≈ 6 dB difference at full modulation depth).

[Figure 8 about here.]

We next investigated whether a similar excess of neural adaptation was observed with responses to alternating spectral cues (i.e. alternating between stimulating electrodes) in CI listeners. AC-FRs generated by alternating electrodes (6.1 Hz; Fig. 9) were clearly observed, although of smaller amplitude than AC-FRs from NH listeners. A LME analysis of the 6.1 Hz response indicated that neither GFP [$F(2, 7.99) = 0.05$, $p = 0.951$, $R_{m/c}^2 = 0.007/0.016$] nor latencies [$F(2, 8) = 1.25$, $p = 0.336$, $R_{m/c}^2 = 0.124/0.307$] were affected by which *electrode pair* was assessed in the alternating paradigm. In general, although relatively small, significant responses were observed at all electrodes; Hotelling-T2 F-values ranged between 3.32 and 24.0 (critical F value of 3.1 for a significance level of $p < 0.05$). A reanalysis of these results without using spatial filtering to reject electrical artefacts led to the same results (data from S4, electrode pair 3 was excluded from this reanalysis due to unacceptable artefact contamination).

The data indicate that AC-FRs from CI listeners were significantly reduced when compared to their transient ACC responses. This reduction is consistent with the hypothesis that, compared

to acoustic stimulation in NH listeners, electrical stimulation in CI listeners generates greater adaptation in cortical neurons in response to speech-like rates alternations of spectro-temporal cues. To quantify the overall level of neural adaptation between low and high alternating rates, the GFP obtained with the ACC and the AC-FR were used to compute the adaptation index (AI) as:

$$AI = 20 \cdot \log_{10} (GFP_{ACC}/GFP_{AC-FR}) \quad (3)$$

[Figure 9 about here.]

First, we assessed whether the AM depth or the electrode discrimination affected the AI in CI. Neither *modulation depth* [$F(1, 5) = 0.95$, $p = 0.374$, $R_{m/c}^2 = 0.057/0.340$ for NH; $F(1, 2) = 0.18$, $p = 0.713$, $R_{m/c}^2 = 0.014/0.614$ for CI] nor *electrode pair* [$F(2, 8) = 2.26$, $p = 0.166$, $R_{m/c}^2 = 0.059 / 0.818$] had a significant effect on the AI. We then assessed whether the AI differed between alternating AM frequencies and electrode stimulation in CI listeners. The analysis indicated that AI was not affected by the *experiment type* [$F(1, 1.78) = 0.272$, $p = 0.659$, $R_{m/c}^2 = 0.004/0.784$]. Next, we compared whether the AI for the AM experiment was different between CI and NH listeners. The mean AI for CI listeners across all conditions was 25.6 dB (s.d. = 6.6 dB), significantly greater [$F(1, 7) = 7.58$, $p = 0.028$, $R_{m/c}^2 = 0.392/0.665$] than that for the NH listeners (mean AI of 19.1 dB; s.d. = 4.8 dB). The difference between NH and CI listeners (pooling the AI across all conditions; Fig. 10) was confirmed by a robust non-parametric analysis of variance [$F(1, 18.72) = 14.01$, $p = 0.001$, $\xi = 0.67$].

Furthermore, to control for a potential bias of our artefact cancellation, in which the AC-FR amplitudes may have resulted smaller when removing DSS components, we repeated the previous analysis using the raw AC-FR, i.e. without artefact removal. The results (not reported here) remained virtually the same. These data confirm that, when presented at a speech-like alternating rate, spectral or temporal cues are considerably more adapted in CI listeners than in NH listeners, potentially accounting for some of the difficulties CI listeners have in encoding dynamic spectral and temporal cues critical to speech processing.

[Figure 10 about here.]

4 Discussion

We assessed the ability of CI users to process dynamic patterns of spectro-temporal modulations, including at rates present in typical speech sounds. By means of the alternating ACC paradigm, we elicited cortical responses to dynamic spectral and temporal changes, generated by alternating the stimulating electrode of a pair of electrodes, or the AM frequency of a single electrode, respectively. We obtained significant cortical responses when alternating between stimulation of pairs of CI electrodes at a slow, 0.5 Hz, rate. For the same, slow alternating rate, we also demonstrated that transient ACC responses were evoked in CI listeners by alternating the AM frequency, from all participants at the high (801 pps) pulse rate, and for some listeners at the low (126 pps) rate.

ACC responses were dramatically reduced in NH listeners when AM depth was reduced from 100% to 50%. This was not the case for CI users, where similar high-amplitude responses were obtained for both 100% and 50% AM depths, when stimulating at the highest pulse rate (801 pps). Overall, anodic (QP-A) pulse trains were easier to discriminate than cathodic (QP-C) pulse trains. However, cortical ACC responses evoked with low-rate pulse trains (126 pps) were weak regardless of stimulus polarity. Finally, we showed that rapid alternating cues evoked AC-FRs in both NH and CI listeners. However, responses from CI participants were much more adapted than those from NH listeners, confirming that cortical activity decreased faster for CI than NH listeners.

4.1 Neural adaptation to changes in an ongoing stimuli

Multiple factors likely contribute to poor speech understanding in CI users. Whilst neural entrainment to speech envelopes plays a critical role in the encoding of speech (e.g. Ding and Simon, 2012; Peelle et al., 2013; Broderick et al., 2019), our data suggest that the inability to represent changes in spectro-temporal cues at rates common to natural speech may be critical in CI listeners. This does not preclude CI users exploiting already degraded static spectro-temporal processing. Although it was not possible to evoke robust AC-FRs to rapid (≈ 7 Hz) alternating rates in CI users, significant ASSRs (20 and 35 Hz) were observed at both slow (126 pps) and fast (801 pps) pulse rates. This suggests that strong neural synchronization to electrical stimulation in primary auditory neurons allows for sub-cortical neural generators of the ASSR to follow AM

frequencies in the range of a few tens of Hz, but that the cortical generator of the AC-FR is rapidly adapted and, consequently, unable to track AM transitions that occur at speech-like rates. This supports the hypothesis that rapid changes in spectro-temporal cues conveyed by highly synchronized electrical pulses induce greater neural adaptation at the cortical level. This is not the case in NH listeners, for whom neural activity is likely less synchronized, allowing for sparse neural activation and neural tracking of rapidly fluctuating cues.

If the ability to encode and discriminate dynamic spectro-temporal cues at a cortical level is intrinsic to speech understanding, it suggests one reason why CI listeners may be limited in exploiting these cues (Zhang et al., 2019); excessive adaptation in the cortex to changing spectro-temporal cues. Together with the limited number of independent channels, the highly synchronized neural response of the inner ear may lead to an overly adapted neural response at the cortical level.

4.2 Temporal processing in NH and CI listeners

At a pulse rate of 801 pps—representative of pulse rates employed in clinical devices—ACC responses were well defined, and we observed two distinct differences between NH and CI listeners. First, in NH listeners, ACC and AC-FRs were significantly smaller when the AM depth was reduced to 50%. Conversely, CI listeners were less sensitive to changes in AM depth—likely arising from the sharp amplitude growth functions and narrow dynamic range typical of CI users (e.g. Cohen, 2009; Zeng and Shannon, 1992; Zeng et al., 2002)—and ACC responses were indistinguishable in terms of magnitude and latency for the two modulation depths (Figures 6K & L), whilst AC-FRs were absent. This indicates that the poor perceptual difference between fully and partially AM'ed stimuli for CI listeners is reflected by large cortical activations, and indicates that shallow AMs may be overrepresented, potentially deteriorating speech perception (Moore et al., 1995; Moore and Glasberg, 1993). At a lower pulse rate (126 pps), cortical representation was poor compared to 801 pps, where ACC responses were well defined and behavioural performance was at ceiling, supporting the hypothesis that low pulse rates have a detrimental effect on the representation of temporal envelopes.

4.3 Effect of polarity in AM encoding

At low pulse rates (126 pps) AM frequencies conveyed by anodic QP-A pulses were easier to discriminate (at both AM depths) than those conveyed by cathodic (QP-C) pulses, consistent with observations in CI listeners that anodic pulses are more effective than cathodic ones (Macherey et al., 2008; Undurraga et al., 2010, 2012, 2013). This supports the hypothesis that interactions between place-of-excitation and pulse polarity may distort envelope cues at the level of the auditory nerve (Joshi et al., 2017). In our study, two participants showed some polarity sensitivity, and of the 5 instances where CI listeners were unable to distinguish behaviourally different AM frequencies, 4 came from these two participants. Contrary to their behavioural scores, however, transient ACC/ASSRs represented the predominant feature of their cortical responses (albeit of relatively small magnitude), but were largely absent in other participants. This suggests that responses from participants with lower anodic thresholds may reflect highly synchronized somatic or post-somatic activation of ANFs, compared to the relatively less-synchronized activation of peripheral processes evoked by cathodic stimulation (Undurraga et al., 2010, 2012, 2013). Furthermore, Macherey et al. (2017) measured loudness growth functions in human CI listeners using both QP-C and QP-A pulse trains presented at a rate of 100 pps. They found that loudness growth functions with QP-A pulse trains increased monotonically with the stimulus amplitude. This was not the case for QP-C pulse trains, and several listeners show clear non-monotonic loudness growth functions. Thus, it may also be possible that the effective modulation depth is shallower for QP-C than for QP-A pulse trains, difficulting the encoding the different AM frequencies, and thus, discrimination.

4.4 AM discrimination as a measure of temporal envelope processing

Our use of an AM frequency discrimination task to assess temporal envelope processing is highly novel. CI users are able to understand speech based on the AM'ed cues conveyed by the envelope at each channel (Shannon et al., 1995; Fu and Shannon, 2002), and many studies have measured AM discrimination thresholds behaviourally (e.g. Galvin and Fu, 2005; Garadat et al., 2012; Chatterjee and Peng, 2008; Galvin and Fu, 2009) or objectively (Han and Dimitrijevic, 2015; Han and Dimitrijevic, 2020; Gransier et al., 2019). However, these studies compared an unmodulated reference signal to a modulated target signal in which the AM depth was reduced until it was

no longer distinguishable from the reference. A problem associated with this method is the confounding factor of available loudness cues, which may contribute to participants' judgments in a detection or discrimination task (McKay and Henshall, 2010). Whilst these cue can be minimized—for example, by adjusting signal intensity to maintain loudness perception as a function of the AM depth, or roving the overall signal amplitude between trials—this procedure usually requires obtaining significant behavioural data and, even then, does not remove the cue as a potential confound. This problem can be largely avoided if AM frequency discrimination is measured with both AM frequencies having the same AM depth—the loudness of low pulse rate AM pulses is almost constant (Kreft et al., 2010; Gransier et al., 2016; McKay and Henshall, 2010). Thus, provided that the AM frequency difference is above the discrimination limen at 100% AM depth, the AM-frequency discrimination task should reflect the participant's ability to exploit temporal envelope cues per se, and not loudness cues.

4.5 Neural processing of spectral and temporal envelope cues

The ACC response is largely dependent on the extent to which a stimulus change can be perceived. A potential explanation for AM alternations generating longer-latency responses than electrode changes in CI listeners is that neural responses to AM were perceptually weaker than those elicited by electrode changes. However, taking into account that behavioural discrimination of AM was at ceiling at 801 pps, an alternate interpretation is that the generator of the response requires a longer integration time to process the change in AM. The average period of a 20 Hz (50 ms) and a 35 Hz (28 ms) AM cycle is 39 ms, almost identical to the latency difference (42 ms) between electrode and envelope ACC responses, suggesting that neurons require at least one cycle of the AM stimuli to process the switch in AM frequency.

4.6 Conclusions

Using the alternating ACC paradigm at rates that regulate the amount of neural adaptation, we demonstrate that spectro-temporal cues are considerable more adapted in CI listeners compared to NH listeners. Excess of cortical adaptation to spectro-temporal modulation patterns may limit the extraction of important dynamic cues for the understanding of speech, in particular, in background noise and may explain some of the variability in performance observed across CI

listeners, in addition to the limited number of independent channels, and differential sensitivity to stimulus polarity reflecting neural health. Finally, whilst the results were consistent across the high-performing CI listeners that took place in our study, the sample size is very limited and results should not be generalized. Further studies should investigate the degree of neural adaptation to speech-like alternating rates in a much larger sample size.

Acknowledgements

Authors acknowledge support from the National Cochlear Implant Users Association of the UK and from the SCIC. The authors also thank Dr WaiKong Lai (SCIC) for his useful comments and discussions during data collection. The authors also thank Simone de Rijk for implanting the human head and assistance during the artefact assessment. Author Deborah Vickers is the recipient and supported by a Medical Research Council Senior Fellowship (MR/2002537/1), author Manohar Bance is supported by Evelyn Trust Cambridge, and authors Jaime Undurraga, Lindsey Van Yper, and David McAlpine are recipients and supported by a grant from the Australian Research Council (FL160100108).

References

- Bierer, JA (2007). “Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration”. *The Journal of the Acoustical Society of America* 121, 1642–1642.
- Bierer, JA, Bierer, SM, and Middlebrooks, JC (2010). “Partial tripolar cochlear implant stimulation: Spread of excitation and forward masking in the inferior colliculus.” *Hearing research*, 1–9.
- Broderick, MP, Anderson, AJ, and Lalor, EC (2019). “Semantic Context Enhances the Early Auditory Encoding of Natural Speech”. *J Neurosci* 39, 7564–7575.
- Brown, CJ, Etler, C, He, S, O’Brien, S, Erenberg, S, Kim, JR, Dhuldhoya, AN, and Abbas, PJ (2008). “The Electrically Evoked Auditory Change Complex: Preliminary Results from Nucleus Cochlear Implant Users”. *Ear Hear* 29, 704–717.
- Busby, PA and Clark, GM (2000). “Electrode Discrimination by Early-Deafened Subjects Using the Cochlear Limited Multiple-Electrode Cochlear Implant”. *Ear and Hearing* 21, 291–304.

- Chandrasekaran, C, Trubanova, A, Stillittano, S, Caplier, A, and Ghazanfar, AA (2009). “The Natural Statistics of Audiovisual Speech”. *PLoS Comput Biol* 5. Ed. by KJ Friston, e1000436.
- Chatterjee, M and Peng, SC (2008). “Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition”. *Hearing Research* 235, 143–156.
- Chen, H and Zeng, FG (2004). “Frequency modulation detection in cochlear implant subjects”. *The Journal of the Acoustical Society of America* 116, 2269–2277.
- Cohen, LT (2009). “Practical model description of peripheral neural excitation in cochlear implant recipients: 1. Growth of loudness and ECAP amplitude with current.” *Hearing research* 247, 87–99.
- Croghan, NBH, Duran, SI, and Smith, ZM (2017). “Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility”. *The Journal of the Acoustical Society of America* 142, EL537–EL543.
- Dajani, HR and Picton, TW (2006). “Human auditory steady-state responses to changes in interaural correlation.” *Hearing research* 219, 85–100.
- De Cheveigné, A and Simon, JZ (2008). “Denoising based on spatial filtering.” *Journal of neuroscience methods* 171, 331–9.
- Deprez, H, Gransier, R, Hofmann, M, van Wieringen, A, Wouters, J, and Moonen, M (2017). “Characterization of cochlear implant artifacts in electrically evoked auditory steady-state responses”. *Biomedical Signal Processing and Control* 31 (Supplement C), 127–138.
- Ding, N and Simon, JZ (2012). “Emergence of neural encoding of auditory objects while listening to competing speakers.” *Proceedings of the National Academy of Sciences of the United States of America* 109, 11854–9.
- Eisen, MD and Franck, KH (2005). “Electrode interaction in pediatric cochlear implant subjects.” *Journal of the Association for Research in Otolaryngology : JARO* 6, 160–70.
- Field, AP and Wilcox, RR (2017). “Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers”. *Behaviour Research and Therapy*. Best Practice Guidelines for Modern Statistical Methods in Applied Clinical Research 98, 19–38.
- Fu, Qj and Shannon, RV (2002). “Frequency mapping in cochlear implants.” *Ear and hearing* 23, 339–48.

- Galvin, JJ and Fu, QJ (2005). “Effects of Stimulation Rate, Mode and Level on Modulation Detection by Cochlear Implant Users”. *J Assoc Res Otolaryngol* 6, 269–279.
- (2009). “Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users”. *Hearing Research* 250, 46–54.
- Garadat, SN, Zwolan, TA, and Pfungst, BE (2012). “Across-site patterns of modulation detection: relation to speech recognition”. *J. Acoust. Soc. Am.* 131, 4030–4041.
- Geier, LL and Norton, SJ (1992). “The effects of limiting the number of Nucleus 22 cochlear implant electrodes programmed on speech perception.” *Ear and hearing* 13, 340–8.
- Golding, M, Dillon, H, Seymour, J, and Carter, L (2009). “The detection of adult cortical auditory evoked potentials (CAEPs) using an automated statistic and visual detection”. *International Journal of Audiology* 48, 833–842.
- Gransier, R, David, W, Guerit, F, Richardson, M, Middlebrooks, JM, Carlyon, RP, and Wouters, J (2019). “Effect of Modulation Depth on the 40-Hz Auditory Steady-state Response and Auditory Change Complex in CI Users and Normal-Hearing Listeners.” In: The 2019 Conference on Implantable Auditory Prostheses (CIAP). The 2019 Conference on Implantable Auditory Prostheses (CIAP), Lake Tahoe, USA.
- Gransier, R, Deprez, H, Hofmann, M, Moonen, M, van Wieringen, A, and Wouters, J (2016). “Auditory steady-state responses in cochlear implant users: Effect of modulation frequency and stimulation artifacts”. *Hearing Research* 335, 149–160.
- Han, JH and Dimitrijevic, A (2020). “Acoustic Change Responses to Amplitude Modulation in Cochlear Implant Users: Relationships to Speech Perception”. *Front. Neurosci.* 14.
- Han, JH and Dimitrijevic, A (2015). “Acoustic change responses to amplitude modulation: a method to quantify cortical temporal processing and hemispheric asymmetry”. *Front. Neurosci.* 9.
- Harrison, XA, Donaldson, L, Correa-Cano, ME, Evans, J, Fisher, DN, Goodwin, CE, Robinson, BS, Hodgson, DJ, and Inger, R (2018). “A brief introduction to mixed effects modelling and multi-model inference in ecology”. *PeerJ* 6.
- Haywood, NR, Undurraga, JA, Marquardt, T, and McAlpine, D (2015). “A Comparison of Two Objective Measures of Binaural Processing: The Interaural Phase Modulation Following Response and the Binaural Interaction Component”. *Trends Hear* 19.

- He, S, Grose, JH, Teagle, HF, and Buchman, CA (2014). “Objective Measures of Electrode Discrimination with Electrically-Evoked Auditory Change Complex and Speech Perception Abilities in Children with Auditory Neuropathy Spectrum Disorder”. *Ear Hear* 35, e63–e74.
- Henry, BA, McKay, CM, McDermott, HJ, and Clark, GM (2000). “The relationship between speech perception and electrode discrimination in cochlear implantees”. *The Journal of the Acoustical Society of America* 108, 1269–1280.
- Herdman, AT, Lins, O, Roon, PV, Stapells, DR, Scherg, M, and Picton, TW (2002). “Intracerebral Sources of Human Auditory Steady-State Responses”. *Brain Topogr* 15, 69–86.
- Hofmann, M and Wouters, J (2010). “Electrically evoked auditory steady state responses in cochlear implant users.” *Journal of the Association for Research in Otolaryngology : JARO* 11, 267–82.
- Holden, LK, Finley, CC, Firszt, JB, Holden, TA, Brenner, C, Potts, LG, Gotter, BD, Vanderhoof, SS, Mispagel, K, Heydebrand, G, and Skinner, MW (2013). “Factors affecting open-set word recognition in adults with cochlear implants”. *Ear Hear* 34, 342–360.
- John, MS and Picton, TW (2000). “Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements.” *Hearing research* 141, 57–79.
- Joshi, SN, Dau, T, and Epp, B (2017). “Effect of Uncertainty in Site of Spike Generation on Envelope Coding in the Electrically Stimulated Auditory Nerve”. ARO.
- Kreft, HA, Nelson, DA, and Oxenham, AJ (2013). “Modulation Frequency Discrimination with Modulated and Unmodulated Interference in Normal Hearing and in Cochlear-Implant Users”. *J Assoc Res Otolaryngol* 14, 591–601.
- Kreft, HA, Oxenham, AJ, and Nelson, DA (2010). “Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users”. *The Journal of the Acoustical Society of America* 127, 656–659.
- Kuznetsova, A, Brockhoff, PB, and Christensen, RHB (2017). “lmerTest Package: Tests in Linear Mixed Effects Models”. *Journal of Statistical Software* 82, 1–26.
- Lewis, JW (2014). “ALAN CRUTTENDEN , Gimson’s Pronunciation of English, 8th edition. London & New York: Routledge, 2014. Pp. xxvi + 381. ISBN: 978-0-415-72174-5 (hbk), ISBN: 978-1-4441-8309-2 (pbk), ISBN: 978-0-203-78496-9 (ebk) Website: www.routledge.com/cw/”. *Journal of the International Phonetic Association* 44, 297–299.

- Luke, R, Van Deun, L, Hofmann, M, van Wieringen, A, and Wouters, J (2015). “Assessing temporal modulation sensitivity using electrically evoked auditory steady state responses”. *Hear. Res.* 324, 37–45.
- Macherey, O, Carlyon, RP, van Wieringen, A, Deeks, JM, and Wouters, J (2008). “Higher sensitivity of human auditory nerve fibers to positive electrical currents.” *Journal of the Association for Research in Otolaryngology : JARO* 9, 241–51.
- Macherey, O and Carlyon, RP (2012). “Place-pitch manipulations with cochlear implants.” *The Journal of the Acoustical Society of America* 131, 2225–36.
- Macherey, O, Carlyon, RP, Chatron, J, and Roman, S (2017). “Effect of Pulse Polarity on Thresholds and on Non-monotonic Loudness Growth in Cochlear Implant Users”. *JARO*, 1–15.
- Mair, P and Wilcox, R (2020). “Robust statistical methods in R using the WRS2 package”. *Behav Res* 52, 464–488.
- Mathew, R, Undurraga, JA, Li, G, Meerton, L, Boyle, P, Shaida, A, Selvadurai, D, Jiang, D, and Vickers, D (2017). “Objective assessment of electrode discrimination with the auditory change complex in adult cochlear implant users”. *Hear Res.*
- Mathew, R, Vickers, D, Boyle, P, Shaida, A, Selvadurai, D, Jiang, D, and Undurraga, J (2018). “Development of electrophysiological and behavioural measures of electrode discrimination in adult cochlear implant users”. *Hearing Research* 367, 74–87.
- McAlpine, D, Haywood, N, Undurraga, JA, and Marquardt, T (2016). “Objective Measures of Neural Processing of Interaural Time Differences”. In: *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Ed. by Pv Dijk, D Başkent, E Gaudrain, Ed Kleine, A Wagner, and C Lanting. *Advances in Experimental Medicine and Biology*. Springer International Publishing, 197–205.
- McKay, CM and Henshall, KR (2010). “Amplitude modulation and loudness in cochlear implantees.” *Journal of the Association for Research in Otolaryngology : JARO* 11, 101–11.
- Middlebrooks, JC and Snyder, RL (2010). “Selective electrical stimulation of the auditory nerve activates a pathway specialized for high temporal acuity.” *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, 1937–46.

- Moore, BCJ and Glasberg, BR (1993). “Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech”. *The Journal of the Acoustical Society of America* 94, 2050–2062.
- Moore, BCJ, Glasberg, BR, and Vickers, DA (1995). “Simulation of the effects of loudness recruitment on the intelligibility of speech in noise”. *British Journal of Audiology* 29, 131–143.
- Nakagawa, S and Schielzeth, H (2013). “A general and simple method for obtaining R² from generalized linear mixed-effects models”. *Methods in Ecology and Evolution* 4, 133–142.
- Pelle, JE and Davis, MH (2012). “Neural Oscillations Carry Speech Rhythm through to Comprehension”. *Front. Psychol.* 3.
- Pelle, JE, Gross, J, and Davis, MH (2013). “Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension”. *Cereb Cortex* 23, 1378–1387.
- Pfingst, BE, Burkholder-Juhász, RA, Zwolan, TA, and Xu, L (2008). “Psychophysical assessment of stimulation sites in auditory prosthesis electrode arrays.” *Hearing research* 242, 172–83.
- Picton, TW, Vajsar, J, Rodriguez, R, and Campbell, KB (1987). “Reliability estimates for steady-state evoked potentials”. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 68, 119–131.
- Picton, TW, John, MS, Dimitrijevic, A, and Purcell, DW (2003). “Human auditory steady-state responses.” *International journal of audiology* 42, 177–219.
- R Development Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: Foundation for Statistical Computing.
- Rodriguez, R, Picton, TW, Linden, D, Hamel, G, and Laframboise, G (1986). “Human auditory steady state responses: effects of intensity and frequency.” *Ear and hearing* 7, 300–13.
- Rosen, S (1992). “Temporal information in speech: acoustic, auditory and linguistic aspects”. *Phil. Trans. R. Soc. Lond. B* 336, 367–373.
- Ross, B, Picton, TW, and Pantev, C (2002). “Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field”. *Hearing Research* 165, 68–84.
- Shannon, RV, Zeng, FG, Kamath, V, Wygonski, J, and Ekelid, M (1995). “Speech recognition with primarily temporal cues.” *Science (New York, N.Y.)* 270, 303–4.
- Skrandies, W (1990). “Global field power and topographic similarity”. *Brain Topogr* 3, 137–141.

- Undurraga, JA, Carlyon, RP, Macherey, O, Wouters, J, and Wieringen, AV (2012). “Spread of excitation varies for different electrical pulse shapes and stimulation modes in cochlear implants”. *Hear. Res.* 290, 21–36.
- Undurraga, JA, Carlyon, RP, Wouters, J, and van Wieringen, A (2013). “The polarity sensitivity of the electrically stimulated human auditory nerve measured at the level of the brainstem.” *JARO* 14, 359–77.
- Undurraga, JA, Haywood, NR, Marquardt, T, and McAlpine, D (2016). “Neural Representation of Interaural Time Differences in Humans—an Objective Measure that Matches Behavioural Performance”. *JARO*, 1–17.
- Undurraga, JA, van Wieringen, A, Carlyon, RP, Macherey, O, and Wouters, J (2010). “Polarity effects on neural responses of the electrically stimulated auditory nerve at different cochlear sites.” *Hear. Res.* 269, 146–161.
- Valderrama, JT, Torre, Adl, and Dun, BV (2018). “An automatic algorithm for blink-artifact suppression based on iterative template matching: application to single channel recording of cortical auditory evoked potentials”. *J. Neural Eng.* 15, 016008.
- Zeng, FG, Grant, G, Niparko, J, Galvin, J, Shannon, RV, Opie, J, and Segel, P (2002). “Speech dynamic range and its effect on cochlear implant performance”. *The Journal of the Acoustical Society of America* 111, 377–377.
- Zeng, FG and Shannon, RV (1992). “Loudness balance between electric and acoustic stimulation.” *Hearing research* 60, 231–5.
- Zhang, F, Underwood, G, McGuire, K, Liang, C, Moore, DR, and Fu, QJ (2019). “Frequency change detection and speech perception in cochlear implant users”. *Hearing Research* 379, 12–20.

5 Appendix

5.1 Cadaveric recordings

In our previous studies (Mathew et al., 2017, 2018) as well as in the present study we removed electrical artefacts by means of interpolation and DSS. However, it may still be possible that some residual artefacts are present and confused with a biological response. Therefore, we assessed the effectiveness of the methods to subtract electrical artefacts used in this study by measuring these artefacts from the head of an implanted human specimen. This experiment was conducted at University of Cambridge and it was approved by the ethical board of Cambridge University (19/NE/0366) and complied with the [Human Tissue Act \(2004\)](#).

5.1.1 Implantation

The cadaveric head of a 74 years old male human was implanted in the right cochlea with a Cochlear electrode array (CI522). The sample was fresh frozen, without any preservatives to preserve the tissue in optimal conditions, and defrosted within 24 hours of the experiment. The CI was fully inserted (round window approach) and the cochlea was flushed prior to implantation using a 1% saline solution in order to remove air bubbles.

The quality of the electrodes was assessed by impedance measures which were $\approx 2\text{ k}\Omega$ for the electrodes tested (15 and 16).

Recordings were conducted using the same settings as described in Methods, with the only difference we used 32 instead of a 64 EEG recording electrodes. Voltage offsets were within $\pm 20\text{ mV}$ and EEG data were referenced to the channel contralateral to the implant (T7).

5.1.2 Stimuli

If not stated explicitly, all stimuli consisted of QP-C pulses with $43\text{ }\mu\text{s}$ phase width and $7\text{ }\mu\text{s}$ interphase gap presented in monopolar mode (MP 1+2). Stimulation level was 180 CU in all recordings. A total of 241 epochs of $\approx 1\text{ s}$ were presented per condition.

5.1.3 Topographic maps at different CI return electrode locations

To characterize the topographic maps as a function of the location of the extracochlear return electrode used with monopolar stimulation, we presented symmetric biphasic pulses with a long phase width at a low rate (35 pps, 400 μ s per phase, and 7 μ s interphase gap) and recorded the electrical artefact of these pulses. The return electrode (MP1) was placed on three different locations: 1) the promontori, 2) the posterior medial of the temporal muscle, and 3) near orbicularis oculi muscle (lateral edge of orbicularis oculi right eye). As expected, the topographic maps changed with the location of the return electrode (Fig. 11), confirming that the using the topological distribution and the known location of the implanted CI are a good indication to identify residual artefacts.

[Figure 11 about here.]

5.1.4 Effectiveness of interpolation and DSS in the removal of artefact in ASSRs

To assess the extent to which artefact removal can be achieved via interpolation and DSS, we recorded artefacts produced by a 40 Hz 100% AM'ed pulse trains (between 0 and 180 CU) using pulse rates between 100 and 902 pps. First, we proceed by interpolating the electrical artefacts starting at two different time points (the interpolation length was fixed to the pulse rate period). In the first case, interpolation was applied between points located at the onset of the electrical pulse—likely contaminated by strong electrical artefacts—and, in the second case, interpolation was applied between points located before the onset of the pulse (one eighth of the respective pulse rate period), where we expected an optimal artefact cancellation. In agreement with previous studies, interpolation was inefficient when the starting point was too close to the electrical pulse (Fig. 12A), but it was optimal when the starting point preceded the electrical pulse for pulse rates between 100 and 400 pps for electrodes located contralaterally to the CI, and in all electrodes for rates below 400 pps (Fig. 12B).

To fully characterize the recovery of ASSRs from contaminated recordings, we simulated and added “neural responses” to the recorded artefacts. This allowed us to quantify the effect of artefact removal on otherwise highly temporally correlated conditions (Fig. 13A & B). The artefact free neural response was set so that large activity was found in frontal electrodes and small activity towards the back of the head, similar to the observed distribution of these

responses when referenced to left or right mastoids (Fig. 12C & Fig. 13A).

The results showed that interpolation works well, i.e. with minimal distortions, for pulse rates below 400 pps across all recording electrodes (cf. Figures 12C & D). As the pulse rate increased from and above 400 pps, artefacts started dominate the recordings obstructing the detection of the neural source. Removing residual artefact components via DSS allowed the recovery of the ASSR across all pulse rates (Fig. 12E & Fig. 13C). Only a minor residual component was observed at 400 pps on the right side and it was caused by the similarity between neural and artefact amplitudes, which made their separation difficult.

[Figure 12 about here.]

[Figure 13 about here.]

To further investigate the relative effect the ASSR phase on the recovered response at 902 pps via DSS, a total of 16 different starting phases equally separated in steps of 22.5° were simulated at the optimal interpolation point (one eighth of the pulse rate period before the onset of the pulse).

The effect of the electrical artefact (Fig. 14A, Artefact panel) on the target neural response (Fig. 14A, Neural panel) resulted in a distorted amplitude-phase response in which the electrical artefact “drags” the neural response (Fig. 14A, Artefact + Neural panel) towards its own direction. By removing the DSS component with the largest energy on the side of the CI device, the amplitude of recovered ASSR was reliably for many of the simulated neural phases across channels (Fig. 14B). However, the phase and amplitude of the recovered responses were distorted. For some specific neural phases, the recovered amplitudes had smaller amplitudes and their phase was biased in orthogonal directions to the artefact phase than those of the target neural source (Fig. 14A, Recovered panel, B, & D). The estimated GFP to each simulated condition (Fig. 14C) indicated that the average difference between the target (neural) ($-16.5 \text{ dB} \pm 0.03 \text{ dB}$) and the recovered GFP ($-21.3 \text{ dB} \pm 4.40 \text{ dB}$) was 4.8 dB.

These results demonstrate that ASSRs can be recovered from highly contaminated recordings at low pulse rates via interpolation and at high pulse rates via interpolation and DSS artefact removal. However, the amplitude may result suppressed and the phase biased when DSS is applied, depending on the specific relationship between the neural and the phase of the artefact.

It should be noted that the recovered topographic map was similar to that of the neural source across all conditions and, therefore, it can be used as an indication for the presence of a genuine neural response.

[Figure 14 about here.]

5.1.5 Effectiveness of interpolation and DSS in the removal of artefacts by AM frequency and electrode alternations

To fully characterize the limitations of our artefact rejection method, we recorded the electrical artefact from the cadaveric specimen using the AM frequency alternating paradigm. The pulse train was presented at 902 pps and the AM frequencies (20 Hz and 40 Hz) were presented at an alternating rate of 1 Hz. As in the previous section, we simulated and added “neural responses”, both ACC and ASSRs, to the recorded data. A template ACC response was imposed after each AM transition. To assess the representation of relative neural contributions, the ACC amplitudes to each AM differed by half (Fig. 15A). We also simulated ASSRs to each AM frequency, both set with the same amplitude (Fig. 15A & D).

As in the previous section, we investigated the relative effect of interpolation and ASSR phase on the recovered response, a total of 16 different neural phases were simulated equally separated in steps of 22.5° , and 8 different equally spaced starting interpolation points (between 0 and half the pulse rate period back from the pulse onset), yielding a total of 128 different simulations. The recovered ACC responses were excellent across all simulations (cf. Figures 15B, E, C & F). The difference between target and recovered amplitudes or GFP (Fig. 16A & B) was small. The target N1-P2 and cN1-cP2 GFP were $3.04 \text{ dB} \pm 0.70 \text{ dB}$ and $-0.89 \text{ dB} \pm 0.20 \text{ dB}$, respectively (3.94 dB difference), whilst the recovered GFP at N1-P2 and cN1-cP2 were $2.68 \text{ dB} \pm 0.46 \text{ dB}$ and $-2.25 \text{ dB} \pm 0.95 \text{ dB}$, respectively (4.93 dB difference). This resulted in relative differences between target and recovered response of $\approx 1 \text{ dB}$ and confirmed the reliability of our artefact removal method for ACC responses in CI listeners (Mathew et al., 2017).

[Figure 15 about here.]

[Figure 16 about here.]

Next, we investigated the extent to which ASSRs to each AM embedded on the response could be reliably recovered. This is illustrated in Fig. 17, where both the amplitude and phase to each AM frequency are shown at Cz. At 20 Hz (Fig. 17A), the electrical artefact (Artefact panel) was stronger than at 40 Hz (Fig. 17B). As in the previous section, it can be observed how the electrical artefact drags the neural response (Neural panel) towards its own direction, distorting the phase and amplitude of the neural response (Artefact + Neural panel). By interpolating and removing the largest DSS component on the side of the CI, the amplitude of the recovered 20 Hz ASSR was reliably recovered for many of the simulated neural phases across channels (Fig. 18A). However, the phase of the recovered responses was distorted and had a direction orthogonal to the artefact (Fig. 17A and Fig. 18C). At 40 Hz, both amplitude and phase (Fig. 18A & C) of the ASSR was reliably recovered across most conditions. Slight differences were only observed when the phase of the neural response was in the direction of that of the electrical artefact (Fig. 17B & Fig. 18A).

The estimated GFP to each simulated condition (Fig. 18B) indicated that at 20 Hz, the average difference between the target (neural) ($-23.3 \text{ dB} \pm 0.42 \text{ dB}$) and the recovered GFP ($-28.8 \text{ dB} \pm 5.76 \text{ dB}$) was 5.5 dB, whilst at 40 Hz, the average difference between the target (neural) ($-23.2 \text{ dB} \pm 0.18 \text{ dB}$) and the recovered GFP ($-26.2 \text{ dB} \pm 2.75 \text{ dB}$) was 3.0 dB.

[Figure 17 about here.]

[Figure 18 about here.]

As with the alternating AM frequency paradigm, we also evaluated the ability to extract ACC responses from artefact recordings to electrode changes at 902 pps. The results were identical, ACC responses were reliably recovered in all conditions.

These results demonstrate that interpolation and DSS based artefact removal method allow for successfully recovery of ACC and ASSRs. However, ASSR amplitude and phase may result distorted (smaller amplitudes and wrong phase) as a result of the strong temporal correlation between the artefact and neural waveforms.

5.1.6 Effectiveness of DSS to recover AC-FR

Finally, we investigated the ability to recover AC-FRs from contaminated recordings. Because we were unable to observe electrical artifacts related to the alternating rate for either electrode

or AM frequency in the cadaveric recordings, we recreated a highly contaminated AC-FR scenario using the data of the alternating AM frequency paradigm (20 and 40 Hz AM frequencies alternating at 1 Hz at 902 pps). Considering that potential residual artefacts (e.g. DC current caused by electrode changes, or accumulation of charge from one AM to another) should result in periodicities that are half the alternating rate, we simulated neural signals oscillating at twice the rate of the recorded AM frequency artefacts, i.e. we recreated a transposed version of the AC-FRs paradigm with neural sources oscillating at 40 and 80 Hz (Fig. 19A). As in the previous sections, we added the simulated steady-state neural data to the contaminated recordings (Fig. 19B). A total of four different interpolation points (between zero and one eighth of the pulse rate period back from the pulse onset) and 16 different neural phases were simulated (64 different simulations).

The results demonstrated that we were able to recover AC-FRs in all conditions (Fig. 19C). Both phase and amplitudes were similar to that of the target neural source (Fig. 20A-E).

[Figure 19 about here.]

The estimated GFP to each simulated condition (Fig. 20D) indicated that at 40 Hz, the average difference between the target (neural) ($-22.3 \text{ dB} \pm 0.005 \text{ dB}$) and the recovered GFP ($-19.6 \text{ dB} \pm 1.68 \text{ dB}$) was -2.7 dB , whilst at 80 Hz, the average difference between the target (neural) ($-22.5 \text{ dB} \pm 0.005 \text{ dB}$) and the recovered GFP ($-21.0 \text{ dB} \pm 1.22 \text{ dB}$) was -1.5 dB .

[Figure 20 about here.]

We also explored the ability to separate DC currents by means of a periodic squared pulse train as a “neural” source. The results were identical, and we could always recover amplitude and phase reliably. Our results demonstrate that AC-FR amplitude and phase can be recovered with minimal distortions. Deviations were observed only on the side of the CI (Fig. 20C), but these occurred when the phase of the artefact was similar to that of the neural response, an unlikely realistic expectation, since neural responses occur at a different timing than electrical artefacts, i.e. they have a different phase.

5.1.7 Summary and conclusions

The results, where realistic artefacts were obtained from an implanted human head, provide strong evidence that combining interpolation and DSS can significantly reduce CI artefacts.

ACC and AC-FRs can always be recovered at high (via interpolation and DSS) or low (via interpolation) pulse rates, whilst ASSRs can be reliably recovered at low pulse rates (via interpolation). At high pulse rates, the recovered ASSR (via interpolation and DSS) had either similar or smaller in amplitudes than the neural source. The results also demonstrated that, regardless amplitude-phase distortions, topographic maps are well preserved. Thus, topographic maps provide valuable information to assess the nature of the recovered response. We also found that ASSR amplitude and phase should be carefully assessed. For example, if the phase of the recovered response were orthogonal to that of the contaminated response, this would suggest that the phase and amplitude have been distorted. Conversely, if the phase of the recovered response is not orthogonal to that of the contaminated response, the phase may have been properly recovered, but the amplitude could be smaller than for the neural source. However, it is very unlikely to expect that the neural response will have the same phase than the electrical artefact due to implicit neural delays associated to a response. In either case, our results indicate that the presence of the ASSR (after artefact removal) is indicative of a true neural response, whilst the lack of it does not proof its absence.

5.2 Artefact removal in CI listeners

5.2.1 Transient ACC responses to AM alternation

An example from one participant is shown in [Fig. 21](#) and it is observed how the interpolation successfully removed the electrical artefacts whilst keeping the overall envelope of the EEG recording. Transient ACC responses obtained with a 0.5 Hz alternating rate are also shown in [Fig. 21](#) for one subject, demonstrating a typical P1-N1-P2 morphology.

[Figure 21 about here.]

5.2.2 AC-FR artefact cancellation to electrode alternations

[Fig. 22](#) illustrates the effect of high rate electrode alternations (6.1 Hz) in the time- and frequency-domain for one participant. The left panel shows the projection of the first DSS component on the sensor space. A large artefact contamination following a zero-phase DC drift for each alternation can be observed. In order to minimize artefact contamination, we removed that component.

When artefacts were not removed (left panel), we observed a large sinusoidal component whose phase was almost zero (relative to a sinusoidal signal) or having an opposite polarity (180 degrees) depending on recording electrode. Note that the topographic map was largely biased towards the side of the implant (left hemisphere for this participant).

The strength of the artefact component was strongly reflected in the frequency-domain (middle panel), which it is shown as a large peak at the frequency of a full cycle (3.05 Hz) and its odd harmonics. Removing the artefact component (right panel) reduced the influence of the 3.05 Hz component and its odd harmonics and we could observe a response to the high alternating rate (6.1 Hz). Now, the topographic map shows a distribution that is similar to those observed with the low alternating rates (see [Fig. 3](#)).

[Figure 22 about here.]

Table 1: AM frequency and electrode discrimination experimental details for each CI participant (indicated by the ID). For the AM frequency discrimination, the stimulation rate (in pps), AM frequencies, AM depths (in percentage), thresholds (THR) and comfortable levels (CLs) (in Current Units, the clinical unit used in Cochlear devices) for QP-A and QP-C pulse shapes, are indicated. The AMs were presented at electrode 16. The 50% modulation depth, in dB, is shown in the AM depths column with 0 dB corresponding to 100% modulation depth and -6 dB corresponding to half of the dynamic range measured in a linear scale (i.e. in μA). For the CI electrode discrimination the rate, electrode pairs, and CLs for electrodes tested are indicated. For all participants and experimental conditions pulses consisted of $43 \mu s$ pulse width, and $7 \mu s$ IPG.

ID	AM discrimination				Electrode discrimination		
	Rate [pps]	AM frequencies [Hz]	AM depths [%]	THR-CL [CU]	Rate [pps]	Electrode Pair	CL [CU]
S1	126	20-35	100 & 50 (-4.7 dB) 100 & 50 (-4.7 dB)	QPA: 94-154 QPC: 94-154	827	1: 14-15 2: 15-16 3: 16-17	14: 122 15: 122 16: 122 17: 122
	801	20-35	100 & N.A.	QPC: 57-144			
S2	126	20-35	100 & 50 (-3.9 dB) 100 & 50 (-4.7 dB)	QPA: 140-185 QPC: 140-194	827	1: 14-15 2: 15-16 3: 16-17	14: 168 15: 171 16: 171 17: 171
	801	20-35	100 & 50 (-4.5 dB)	QPC: 123-174			
S3	126	20-35	100 & 50 (-7.3 dB) 100 & 50 (-7.8 dB)	QPA: 107-200 QPC: 107-206	827	1: 14-15 2: 15-16 3: 16-17	14: 165 15: 165 16: 165 17: 165
	801	20-35	100 & 50 (-8.2 dB)	QPC: 72-177			
S4	126	20-35	100 & 50 (-5.4 dB) 100 & 50 (-5.4 dB)	QPA: 130-199 QPC: 130-199	827	1: 14-15 2: 15-16 3: 21-22	14: 166 15: 171 16: 171 21: 185 22: 204
	801	20-35	100 & 50 (-7.1 dB)	QPC: 117-207			
S5	126	20-28	100 & 50 (-4.7 dB) 100 & 50 (-4.7 dB)	QPA: 113-173 QPC: 113-173	827	1: 14-15 2: 15-16 3: 16-17	14: 172 15: 172 16: 169 17: 163
	801	N.A.	N.A.	N.A.			

Table 2: Summary of parameters used for the assessment of spectral and temporal processing in CI and NH listeners. Pulse rate, AM frequencies, AM depths, pulse shape, electrodes, carrier frequency (NH) and alternating rates are listed for each measure performed. The precise pulse rates and modulation frequencies are rounded for simplicity. For one participant, the most apical electrode pair consisted of electrodes 21 and 22. † For one participant, the AM frequency alternated between 20 and 28 Hz. Individual details for CI listeners are provided in [Table 1](#).

Spectral processing						
	Rate [pps]	Electrode Pair		Pulse shape	Alternating rate [Hz] (ACC & AC-FR)	
CI	827	1: 14-15		QP-C	0.5 & 6.1	
	827	2: 15-16		QP-C	0.5 & 6.1	
	827	3: 16-17*		QP-C	0.5 & 6.1	
Temporal processing						
	Rate [pps]	Electrode	AM frequencies [Hz] (ACC & AC-FR)	AM depths [%]	Pulse shape	Alternating rate [Hz] (ACC & AC-FR)
	126	16	20-35 & 20-35 †	100 & 50	QP-A	0.5 & 6.9
	126	16	20-35 & 20-35	100 & 50	QP-C	0.5 & 6.9
	801	16	20-35 & 20-35	100 & 50	QP-C	0.5 & 6.9
Temporal processing						
NH	Carrier [Hz]		AM frequencies [Hz] (ACC & AC-FR)	AM depths [%]		Alternating rate [Hz] (ACC & AC-FR)
	500		20-35 & 19.2-32	100 & 50		0.5 & 6.4

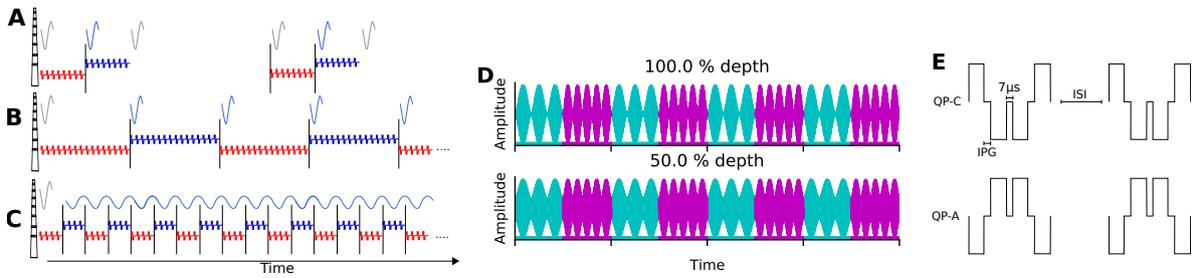


Figure 1: Schematic of stimulation paradigm to discriminate between two electrodes for **A** standard ACC, **B** continuous alternating ACC, and **C** the steady-state AC-FR. Thin gray lines show onset and offset evoked responses associated with start and end of stimulation. Blue waveform illustrates cortical response (ACC and AC-FR) to the transitions. **D** Alternating AM stimuli showing (top) fully modulated (100%) stimulus and (bottom) half-wave modulated (50%) stimulus. Different AM frequencies are shown by the coloured lines. **E** CI stimuli pulse shape, on top, QP-C and bottom QP-A separated by a given ISI.

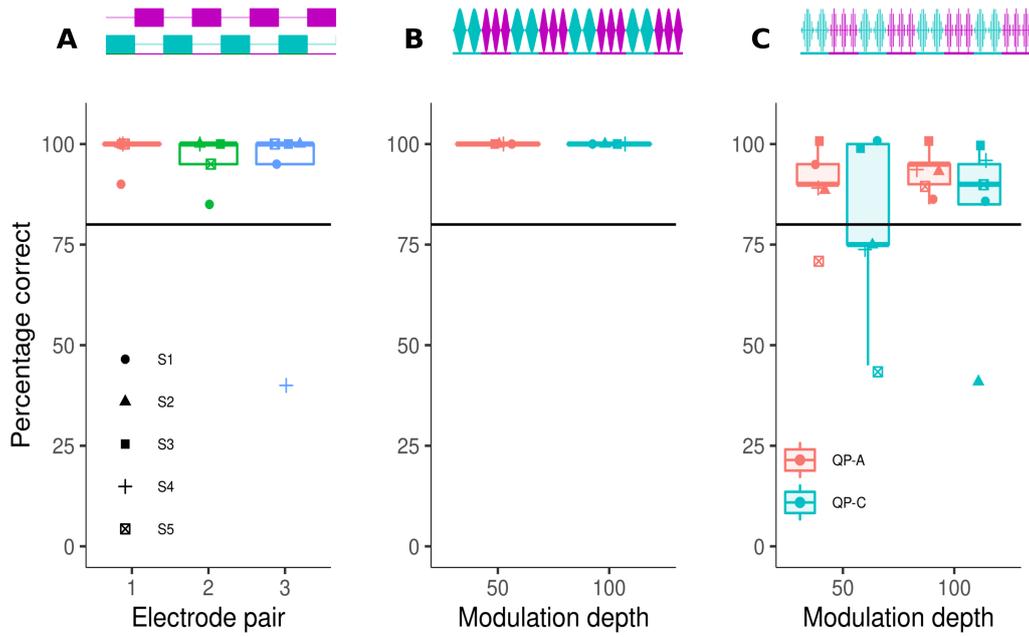


Figure 2: Percentage of correct identification of target stimuli from all CI participants. **A** shows the electrode discrimination for the three adjacent electrode pairs (1: electrodes 14 and 15; 2: electrodes 15 and 16; and 3: electrodes 16 and 17 or 21-22; details in [Table 1](#)). **B** shows amplitude modulation frequency discrimination at two modulation depths using a 801 pps QP-C pulse train. **C** shows the discrimination of the amplitude modulation frequencies at two modulation depths using a 126 pps carrier and two polarities, anodic (QP-A) and cathodic (QP-C). The horizontal line denotes the 80% discrimination threshold. Participants are indicated by each marker shape. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median.

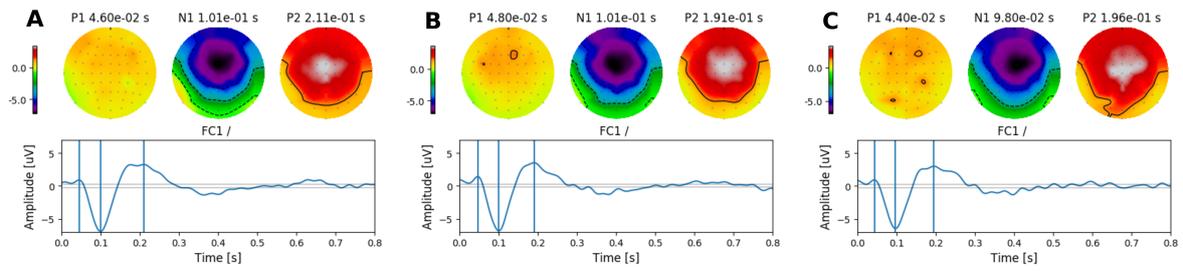


Figure 3: Example of transient ACC responses to electrode changes from participant S5 on the three different electrode pairs. **A** Electrodes 14-15, **B** electrodes 15-16, and **C** electrodes 16-17 obtained at 0.5 Hz alternating rate. P1-N1-P2 peaks are shown in each plot by the vertical lines, respectively. Horizontal grey lines show the residual noise of the recording. Topographic maps at the time of each peak (vertical lines) are shown in each panel. Amplitudes at each channel are indicated by colour code (in μV).

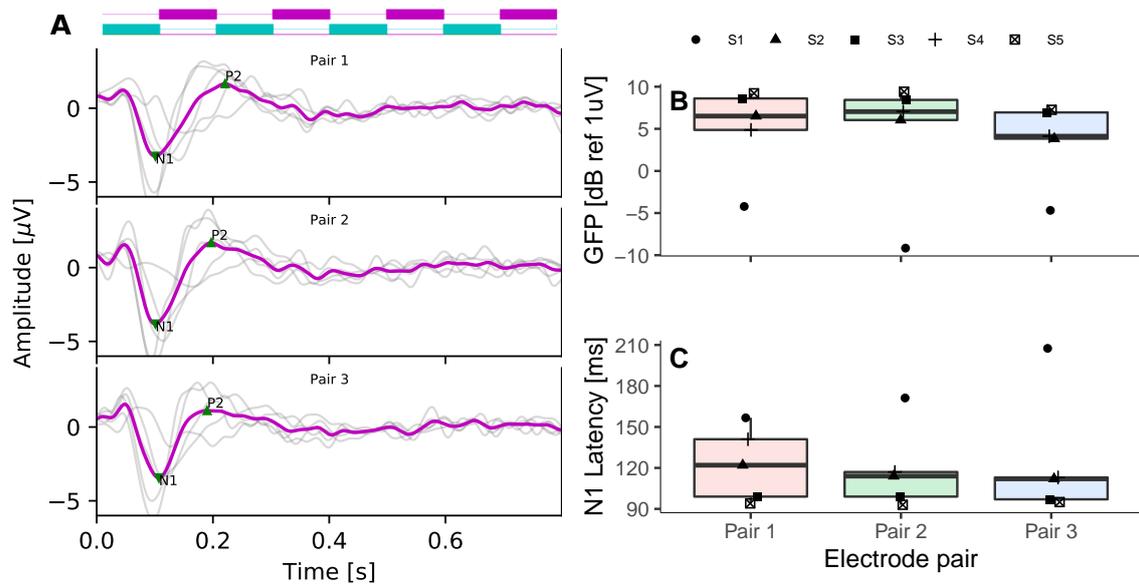


Figure 4: **A** Transient ACC responses to electrode changes at 0.5 Hz alternating rate. Individual responses from the EEG channel with the largest SNR are shown in grey and the group average in magenta. **B** GFP (in dB ref 1 μV) and **C** latencies for all electrode pairs (1: electrodes 14 and 15; 2: electrodes 15 and 16; and 3: electrodes 16 and 17 or 21-22; details in [Table 1](#)). Individual responses are shown by each marker. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median.

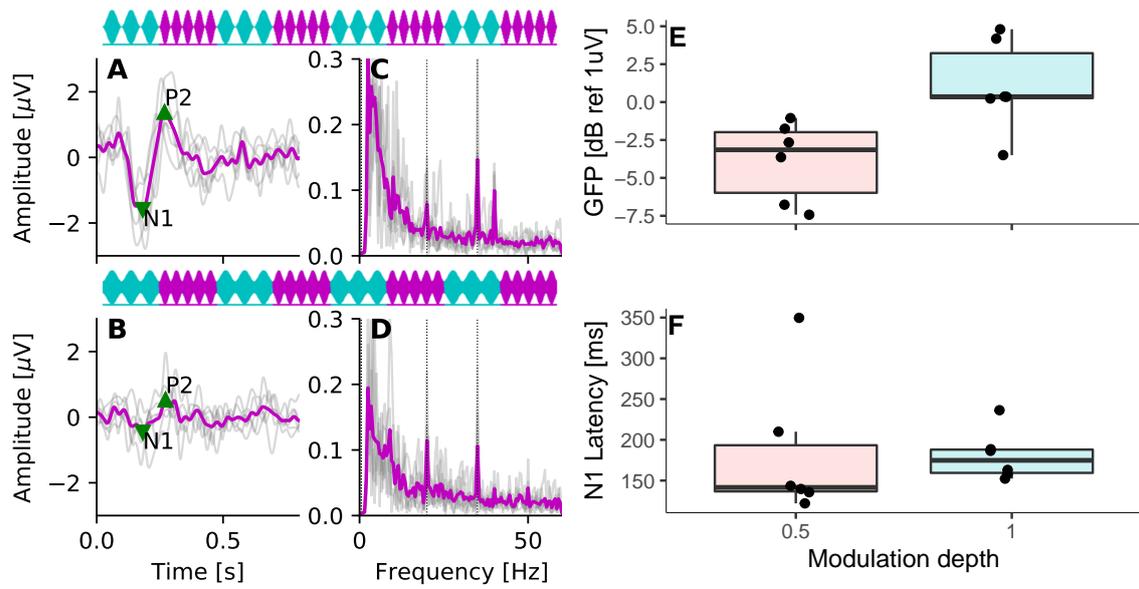


Figure 5: Transient ACCs evoked by continuously alternating at 0.5 Hz the AM frequency between 20 and 35 Hz in NH listeners. **A** and **B** time-domain, **C** and **D** frequency-domain responses at 100% and 50% modulation depth, respectively. Individual responses are shown in grey and average response in magenta. ASSRs evoked by the modulation frequencies are indicated by the dotted vertical lines at 20 and 35 Hz, respectively. **E** GFP (in dB ref 1 μ V) and **F** latencies for both modulation depths. Individual responses are shown by filled dots. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median.

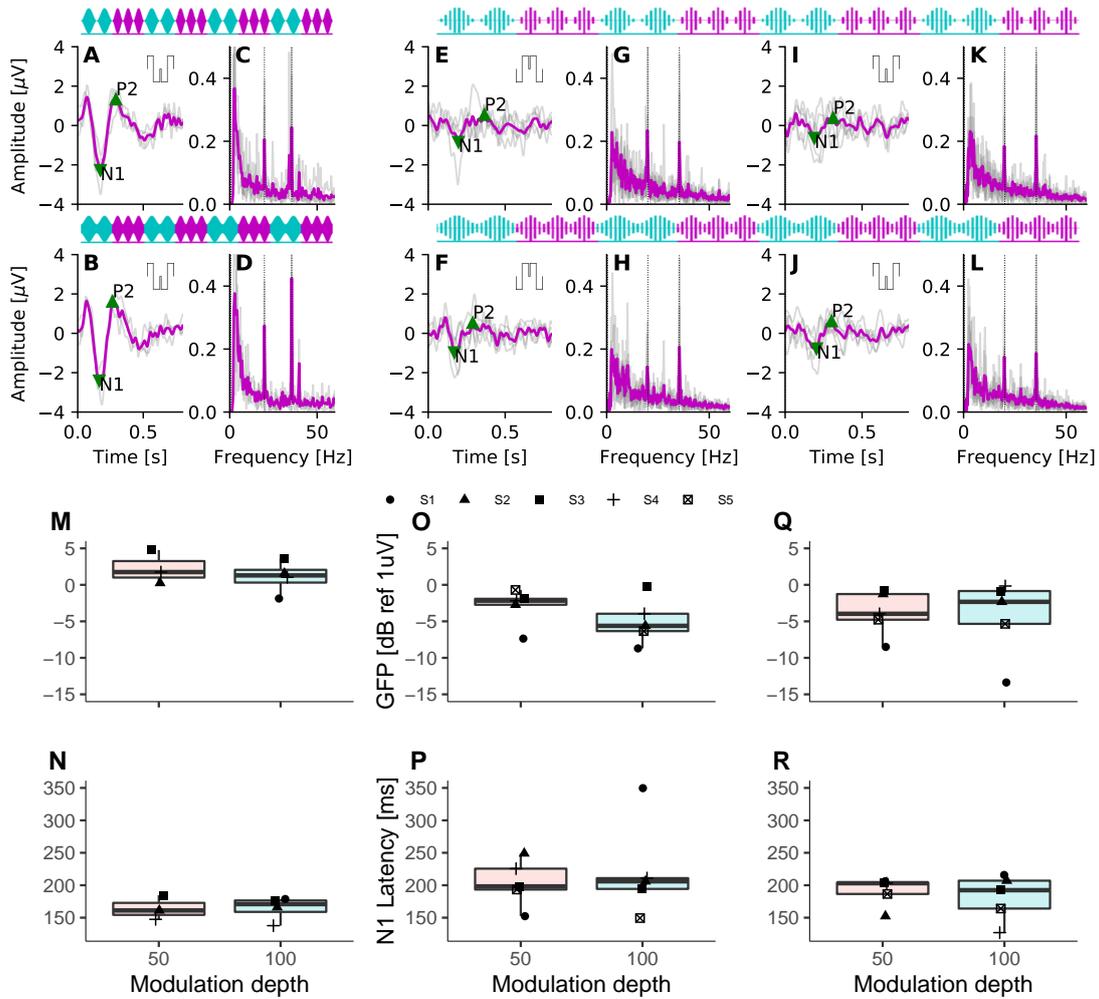


Figure 6: Transient ACC evoked by continuously alternating (0.5 Hz) the AM frequency using pulse trains presented at 801 pps A-D, and 126 pps E-L. **A** and **B** time-domain, **C** and **D** frequency-domain responses at 100% and 50% AM depth using QP-C 801 pps pulse trains. **E** and **F** time-domain, **G** and **H** frequency-domain responses at 100% and 50% AM depth using QP-A 126 pps pulse trains. **I** and **J** time-domain, **K** and **L** frequency-domain responses at 100% and 50% AM depth using QP-C 126 pps pulse trains, respectively. Individual responses are shown in grey whilst average response in magenta. ASSRs evoked by the AM frequencies embedded on the ACC waveform are indicated by the vertical dotted lines at 20 and 35 Hz in the frequency-domain panels. The respective GFP (in dB ref 1 μ V) and latencies are shown in **M** and **N** for QP-C at a pulse rate of 801 pps, **O** and **P** for QP-A at a pulse rate of 126 pps, and **Q** and **R** for QP-C at a pulse rate of 126 pps, respectively. The modulation depth is indicated in the horizontal axis. Individual responses are shown by filled dots. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median.

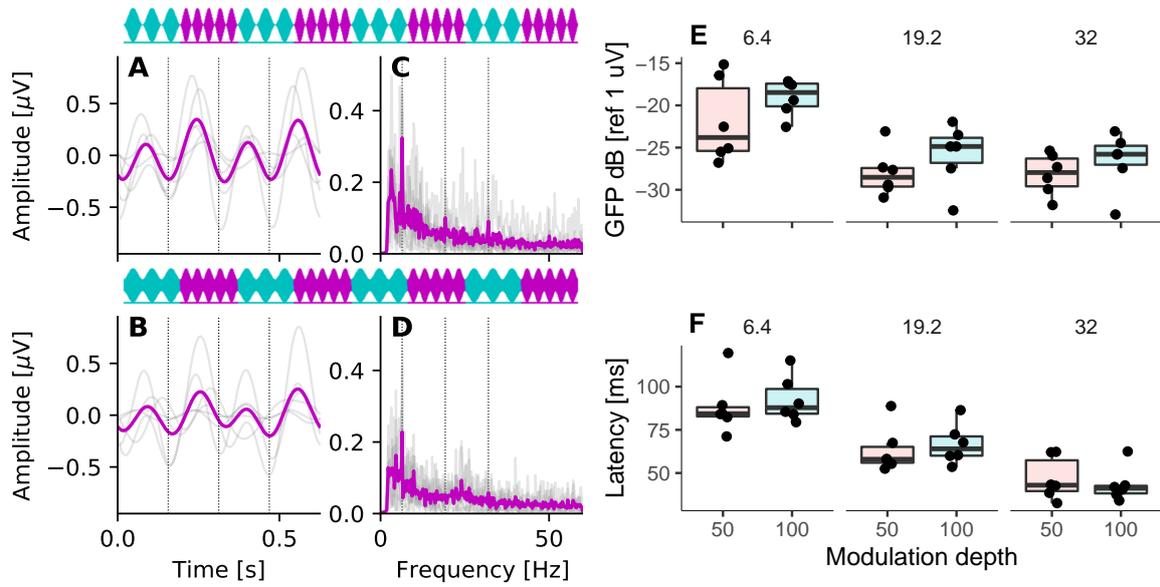


Figure 7: NH listeners AC-FRs evoked by rapid AM alternating rates (6.4Hz). **A** and **B** time-domain, **C** and **D** frequency-domain responses at 100% and 50% AM depth, respectively. Individual responses are shown in grey and average response in magenta. Frequencies corresponding to the alternating rate (6.4Hz) and each AM (19.2 and 32 Hz) are indicated by the vertical dotted lines, in the frequency-domain panels. Vertical lines in time-domain panels indicate the moment of the AM transition. **E** GFP amplitudes and **F** latencies for the alternating rate (6.4Hz), and the AM frequencies (19.2 and 32 Hz) as a function of the modulation depth.

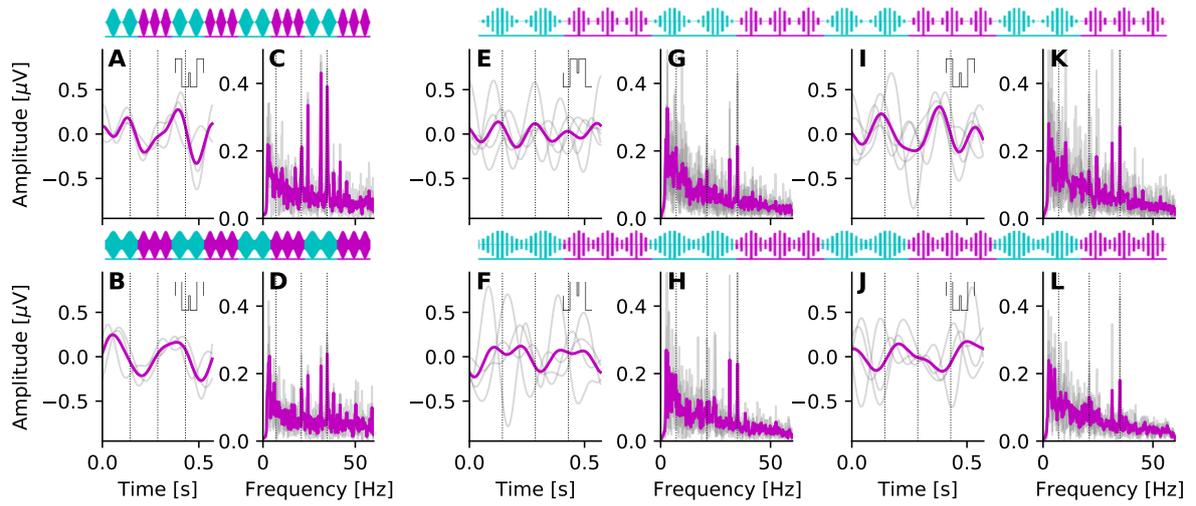


Figure 8: AC-FRs evoked by continuously alternating (6.9 Hz) the AM frequency in CI users. **A** and **B** time-domain, **C** and **D** frequency-domain responses at 100% and 50% AM depth using QP-C 801 pps pulse trains. **E** and **F** time-domain, **G** and **H** frequency-domain responses at 100% and 50% AM depth using QP-A 126 pps pulse trains. **I** and **J** time-domain, **K** and **L** frequency-domain responses at 100% and 50% AM depth using QP-C 126 pps pulse trains, respectively. Individual responses are shown in grey whilst average response in magenta. AC-FRs and ASSRs evoked by the AM frequencies are indicated by the vertical dotted lines at 6.9, 20, and 35 Hz in the frequency-domain panels. Vertical lines in time-domain panels indicate the moment of the AM transition.

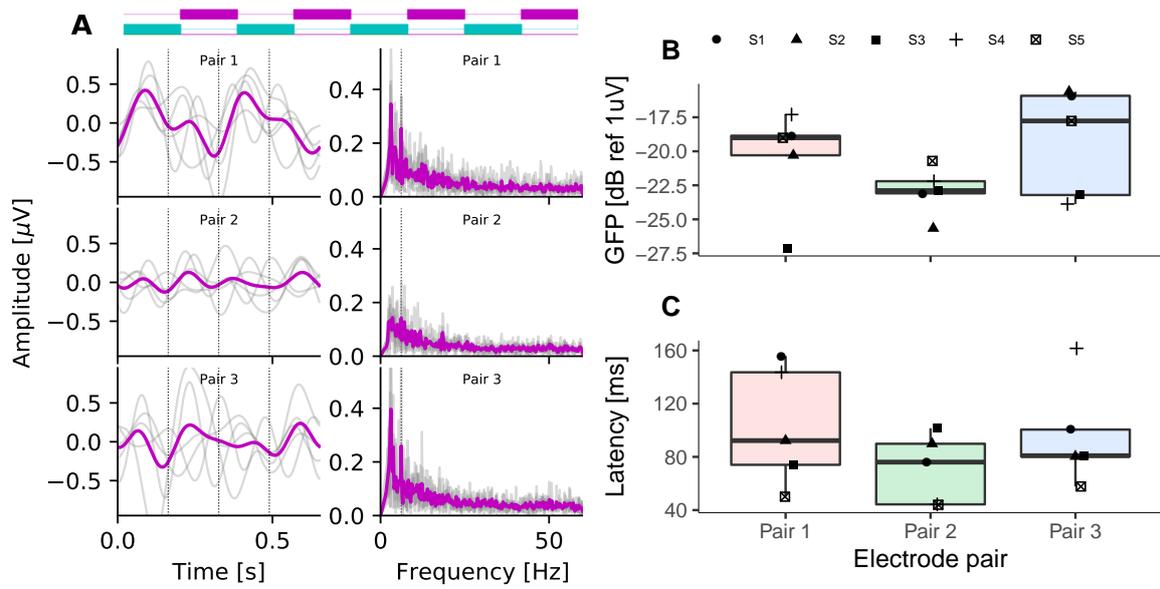


Figure 9: **A** Time and frequency domain AC-FRs obtained at 6.1 Hz alternating rate after artefact removal. Individual responses are shown in grey whilst average response in magenta. Vertical lines in time-domain panels indicate the moment of the electrode transition. The vertical line at 6.1 Hz in frequency-domain responses indicates the alternating rate. **B** GFP (in dB ref 1 μV) and **C** latencies for all electrode pairs. Individual responses are shown by each marker. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median.

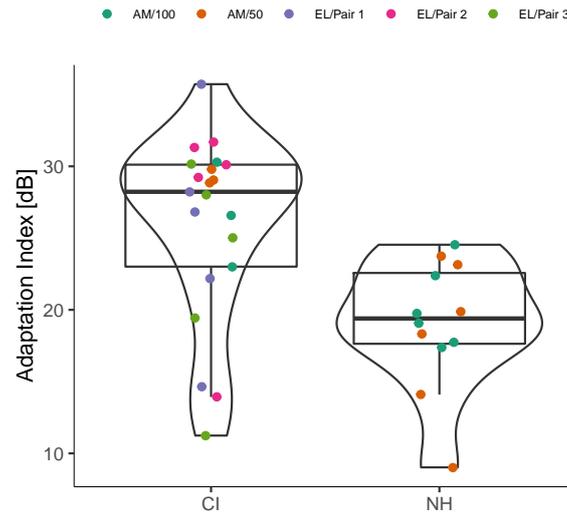


Figure 10: Adaptation index for CI and NH listeners. The source of the adaptation index, i.e., CI or NH listener, is indicated in the horizontal axis. The AM depth or electrode pair are indicated by each color. The boxes show the 25th to the 75th percentile range and the black horizontal shows the median. Violin plots illustrate the full distribution of the data.

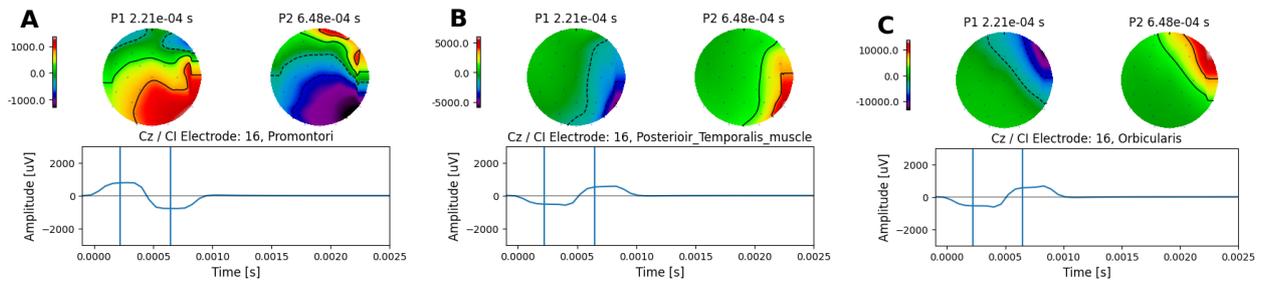


Figure 11: Topographic maps at different reference electrode (MP1) locations. **A** On the promontori, **B** posterior medial of the temporal muscle, and **C** near orbicularis oculi muscle. Topographic maps were obtained for two time points, at the middle of each pulse phase, shown by vertical lines alongside the waveforms.

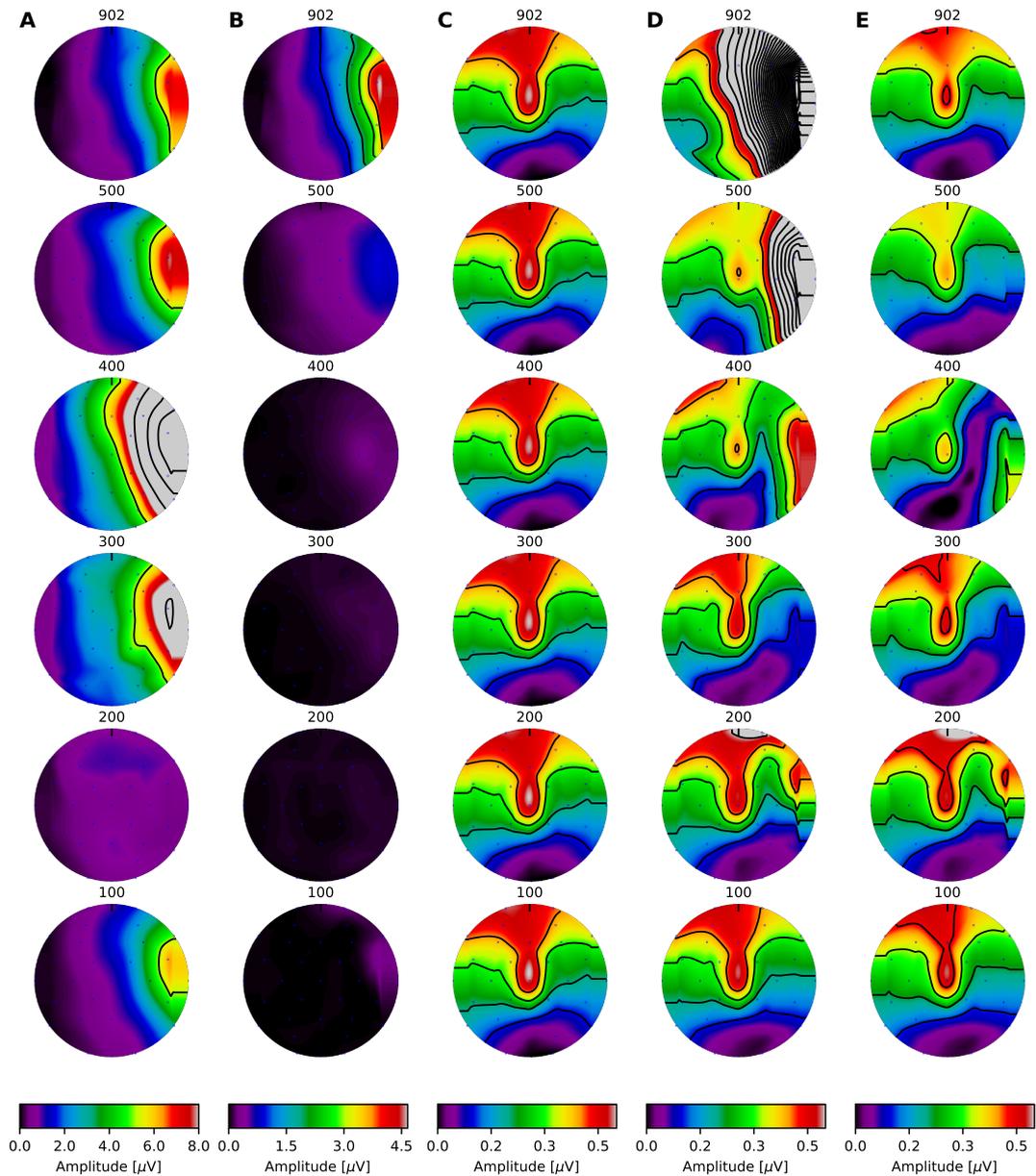


Figure 12: Topographic maps for ASSR amplitudes at different pulse rates and interpolation time points. **A** Electrical artefacts obtained by interpolating between pulses from the onset of each electrical pulse (strong artefact contamination), and **B** interpolating from one eighth of the pulse rate period, before the electrical pulse onset (minimum artefact contamination). **C** Amplitudes of simulated ASSR; **D** topographic map of contaminated response (**B** + **C**); and **E**, topographic maps of recovered neural response. The amplitude of the 40 Hz frequency component at each recording electrode is colour coded and indicated by the colorbar in each panel. Note that the maximum value shown in the colour scales was limited to facilitate visualisation across conditions, however, residual artefacts exceeded 24 μV in many cases.

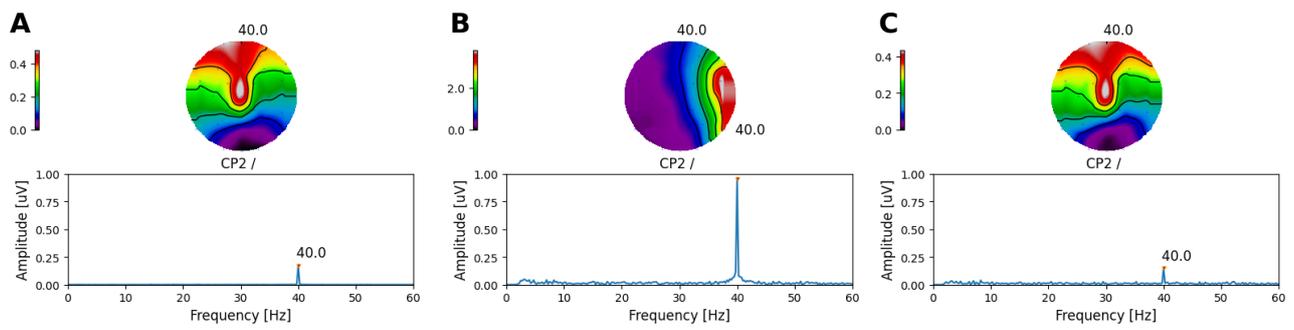


Figure 13: Topographic maps for ASSR at 902 pps. **A** Simulated neural response, **B** recorded electrical artefact plus simulated neural response, and **C** recovered neural response. Topographic maps were obtained for the ASSR frequency response (40 Hz). The response amplitude at each recording electrode is colour coded and indicated by the colorbar in each panel.

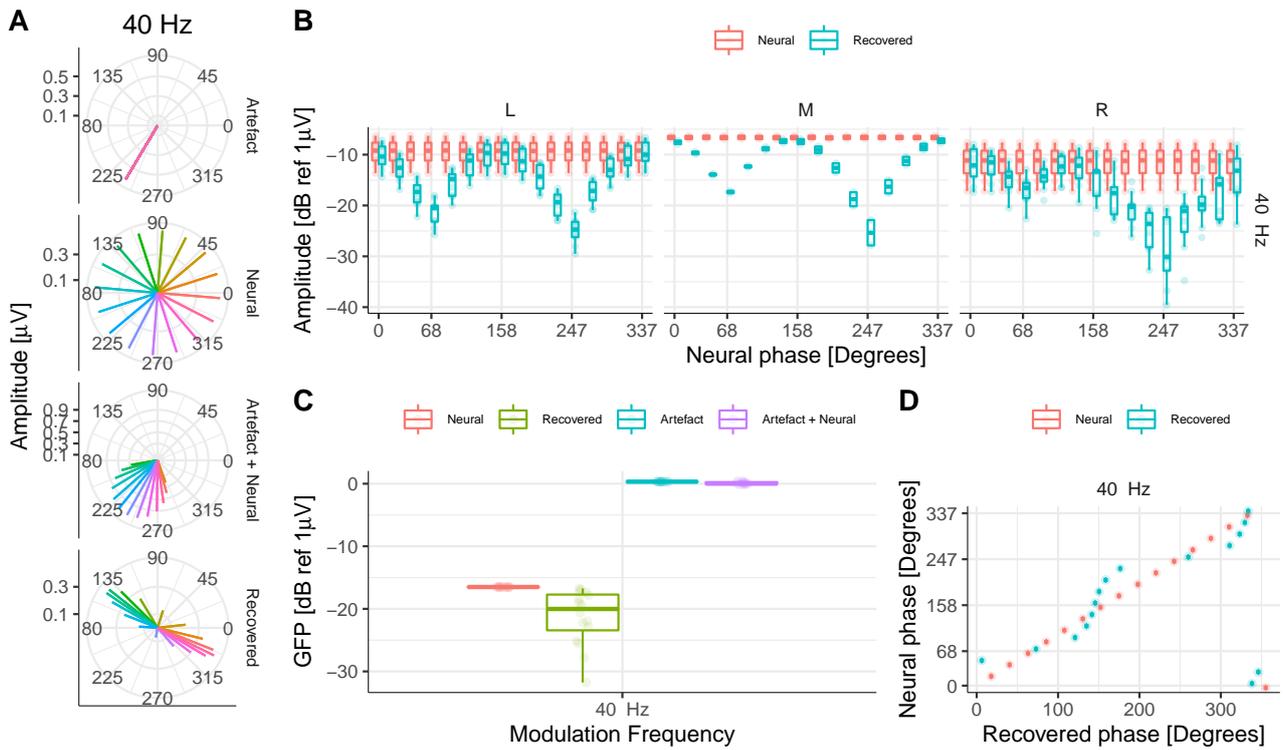


Figure 14: **A** Amplitude and phase of the 40 Hz ASSRs at Cz. Each row shows the amplitude and phase of: artefact only, target (neural), artefact plus neural, and recovered ASSRs, respectively. The different phases of the neural source are colour coded. **B** ASSR amplitudes for central and frontal recording electrodes on the left (L), middle (M), and right (R) side of the head. **C** GFP for all simulated ASSRs. Target (neural), recovered, artefact only, and artefact plus neural data are colour coded. **D** Target (neural) and recovered phase (in degrees) for simulated ASSRs at Cz.

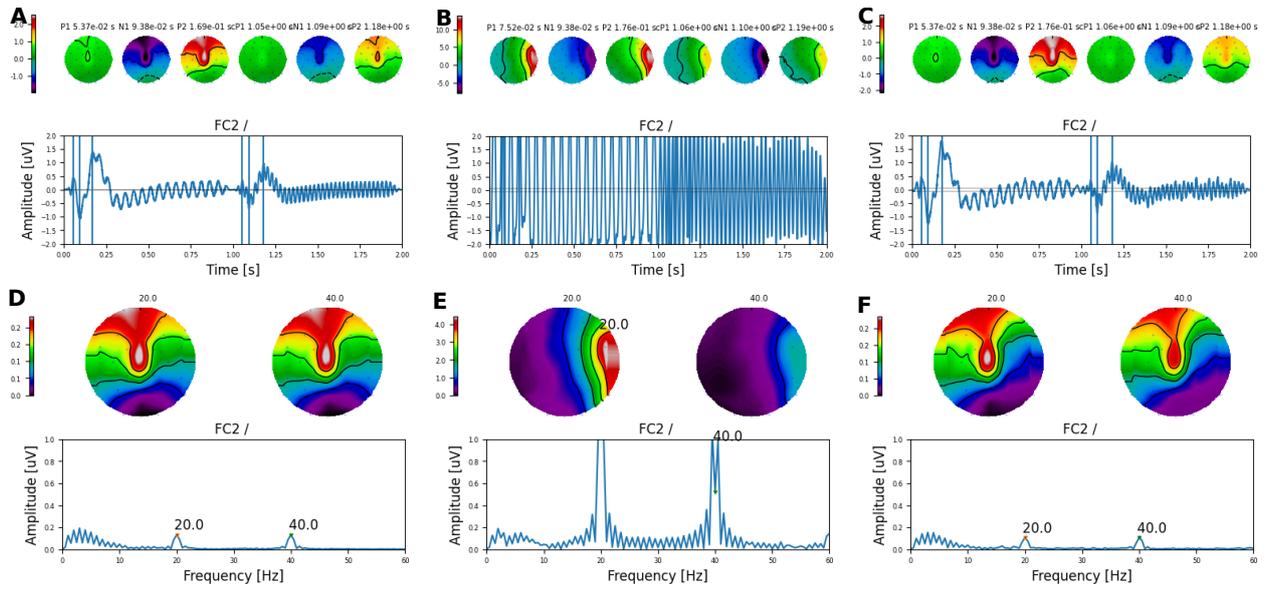


Figure 15: Topographic maps for AM frequency alternating simulations. **A** Target neural response, **B** artefact plus neural response, and **C** recovered neural response. Topographic maps were obtained for the time points shown by vertical lines alongside the waveforms. **D**, **E**, and **F** same as before but in the frequency-domain.

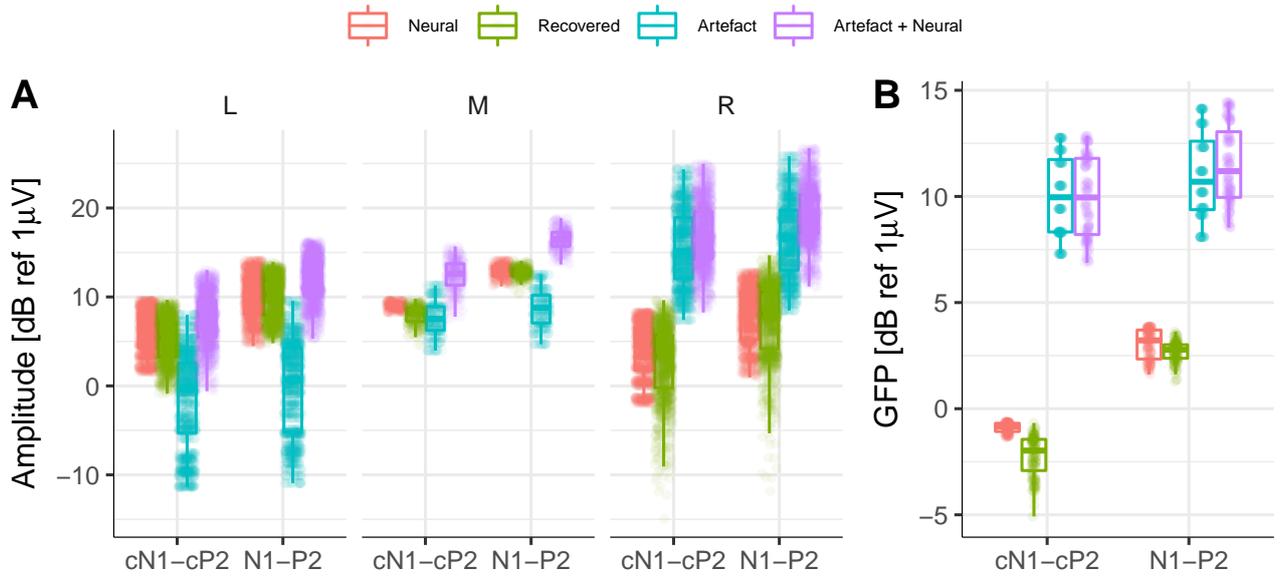


Figure 16: Amplitude and GFP across all simulated conditions. **A** ACC amplitudes for the first (N1-P2) and second (cN1-cP2) simulated responses for central and frontal recording electrodes on the left (L), middle (M), and right (R) side of the head. **B** GFP for the first (N1-P2) and second (cN1-cP2) simulated ACC responses. Target (neural), recovered, artefact only, and artefact plus neural data are colour coded.

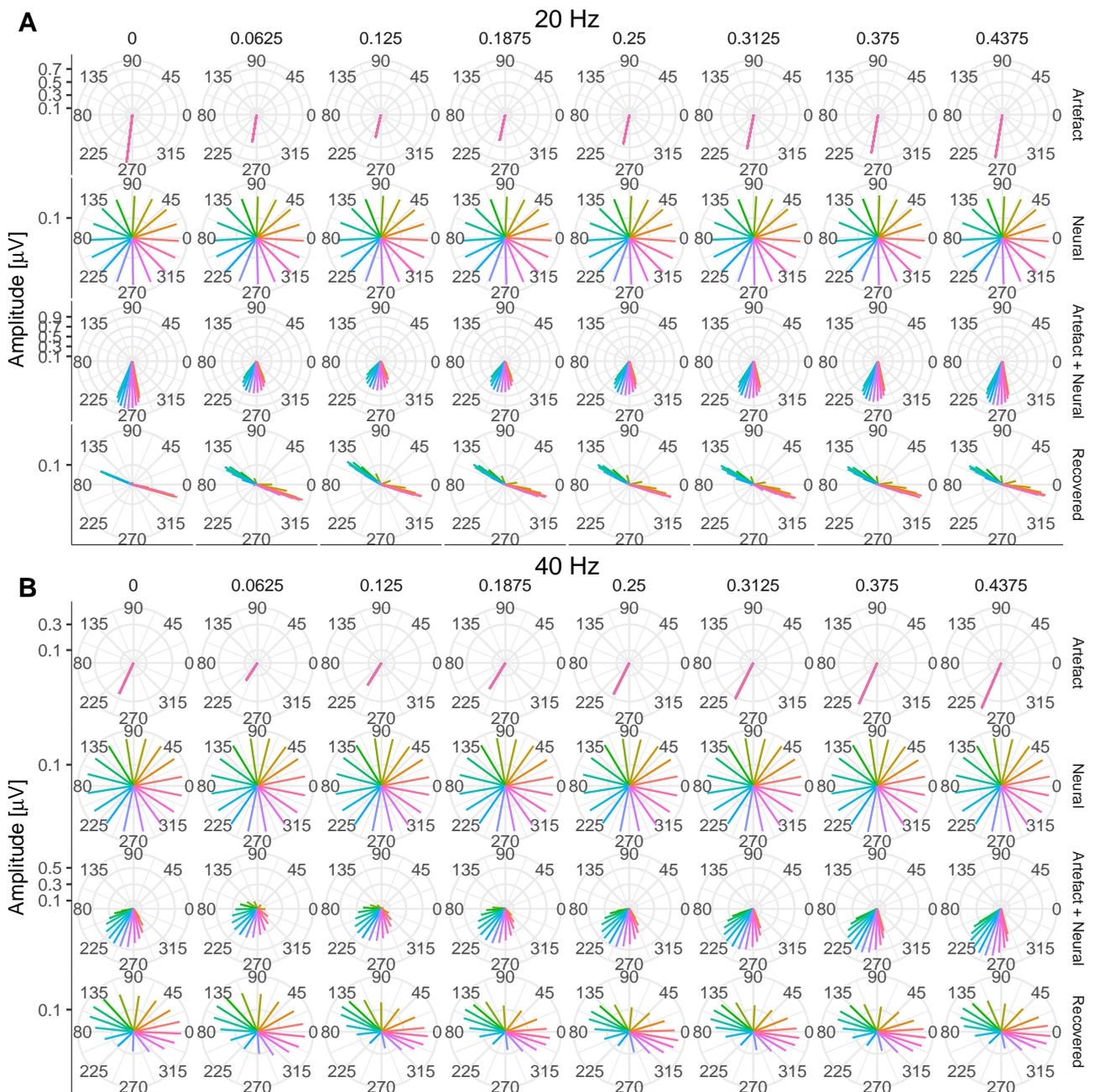


Figure 17: Amplitude and phase of ASSRs to each AM frequency at Cz. **A** for 20 and **B** for 40 Hz. Each column corresponds to a different starting interpolation point (as a fraction of the pulse rate period). Each row shows the amplitude and phase of: artefact only, target (neural), artefact plus neural, and recovered ASSR, respectively. The different phases of the neural source are colour coded.

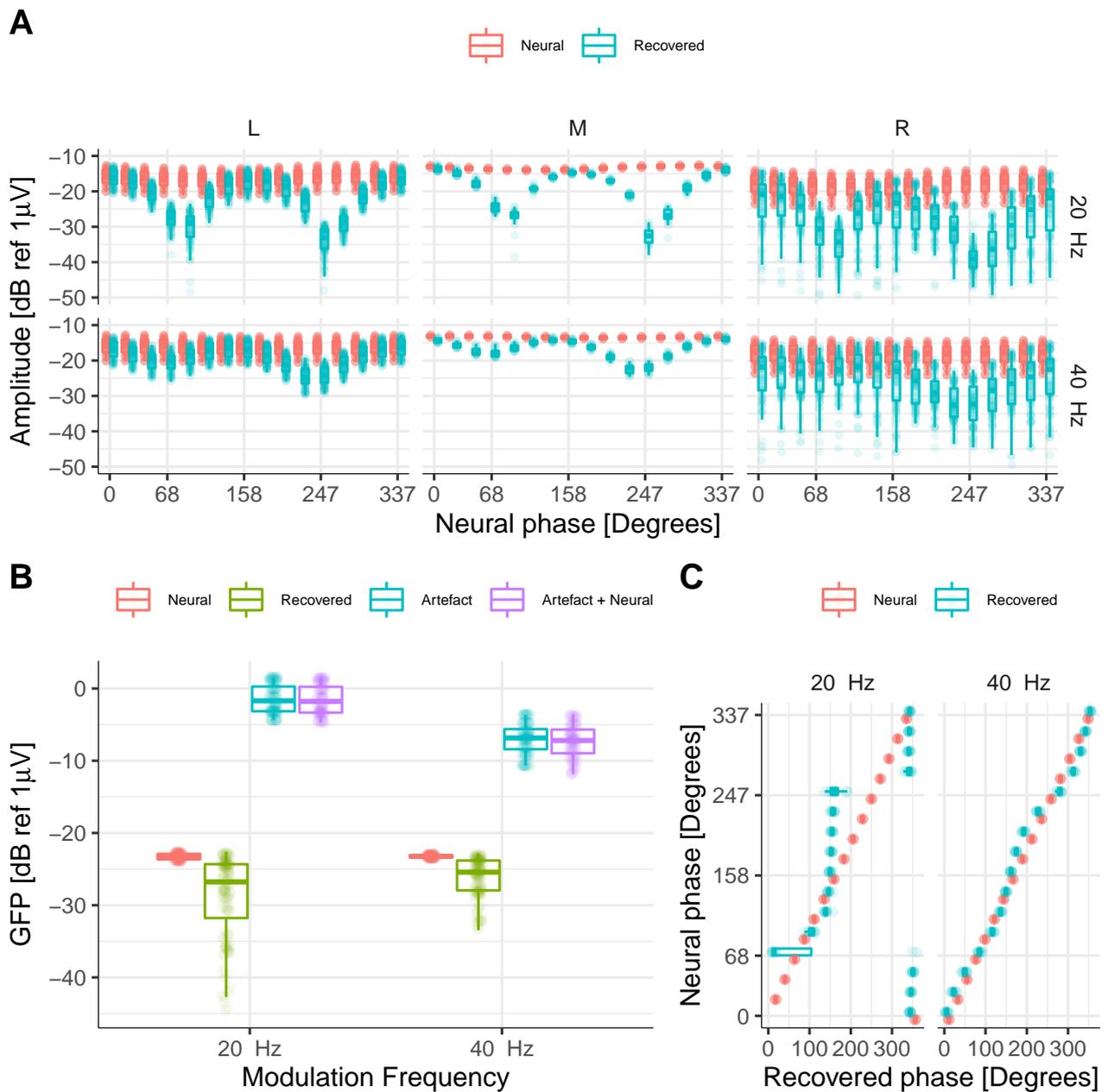


Figure 18: **A** Amplitudes across all frontal and central electrodes for target (neural), and recovered response. **B** GFP across all simulated conditions. Individual points correspond to different interpolation starting points and different neural phases. Target (neural), recovered response, artefact only, and artefact plus neural data are colour coded. **C** Target (neural) and recovered phase (in degrees) for simulated ASSRs to each AM frequency (indicated above each panel) at Cz. Individual points correspond to different interpolation starting points.

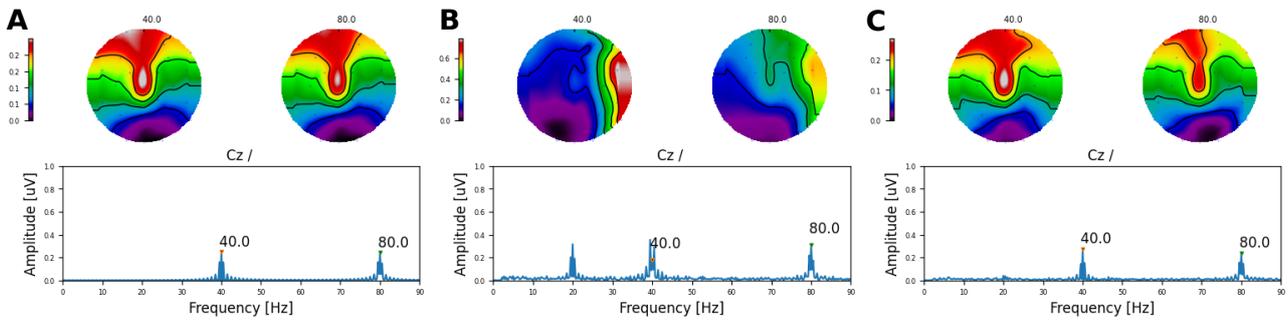


Figure 19: Topographic maps for AC-FR simulations at 902 pps. **A** Target neural response, **B** artefact plus neural response, and **C** recovered neural response. The amplitude is colour coded (in μV). Note that comodulations around the target frequencies are caused by the beating of each frequency (1 Hz) within the analysis window used to compute the frequency-domain response (≈ 4 seconds).

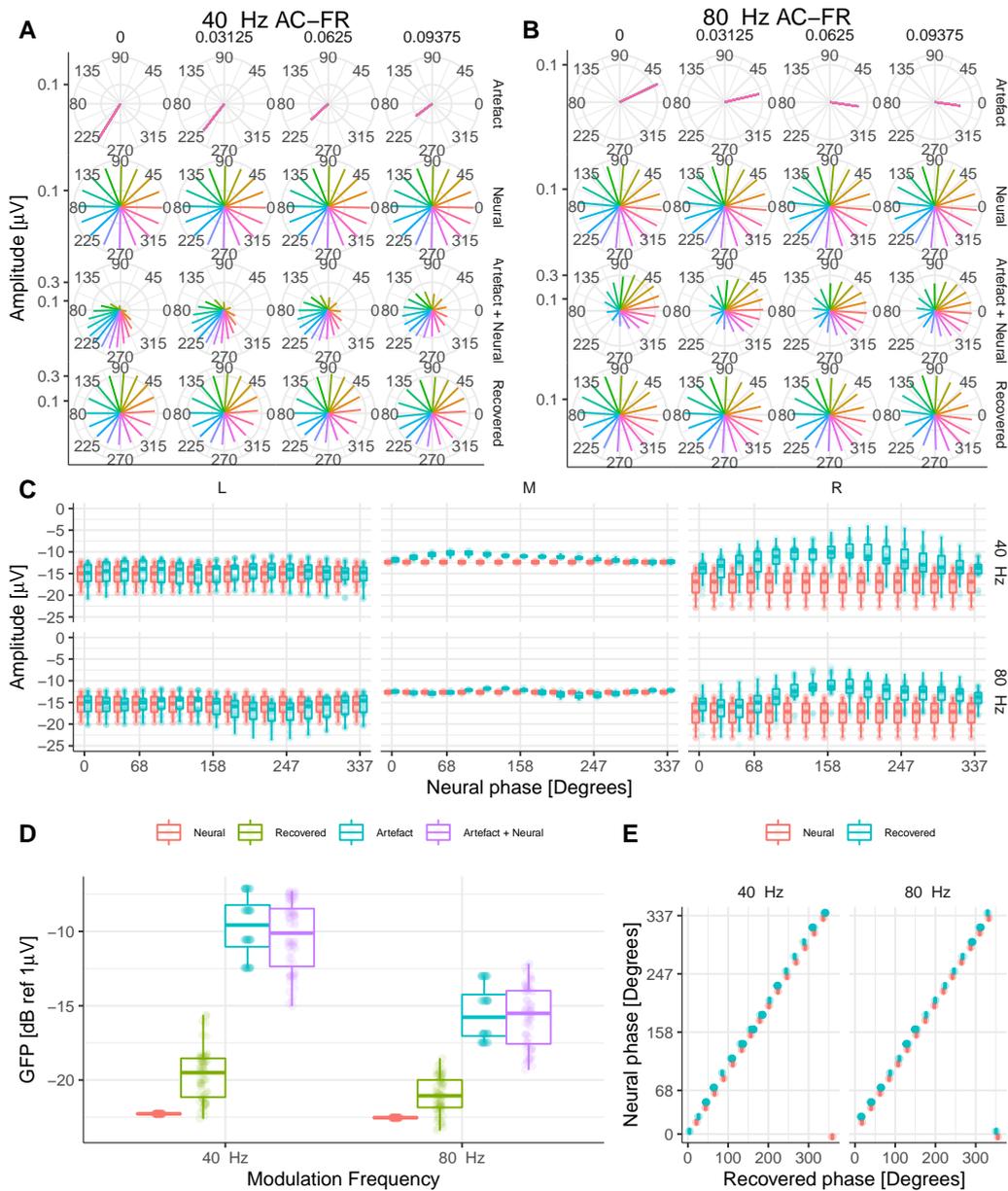


Figure 20: **A** and **B** Amplitude and phase of 40 Hz and 80 Hz AC-FRs at Cz, respectively. Each row shows the amplitude and phase of: artefact only, target (neural), artefact plus neural, and recovered ASSRs, respectively. The different phases of the neural source are colour coded. **C** ASSR amplitudes for central and frontal recording electrodes on the left (L), middle (M), and right (R) side of the head. **D** GFP for all simulated ASSRs. Target (neural), recovered, artefact only, and artefact plus neural data are colour coded. **E** Target (neural) and recovered phase (in degrees) for simulated ASSRs at Cz.

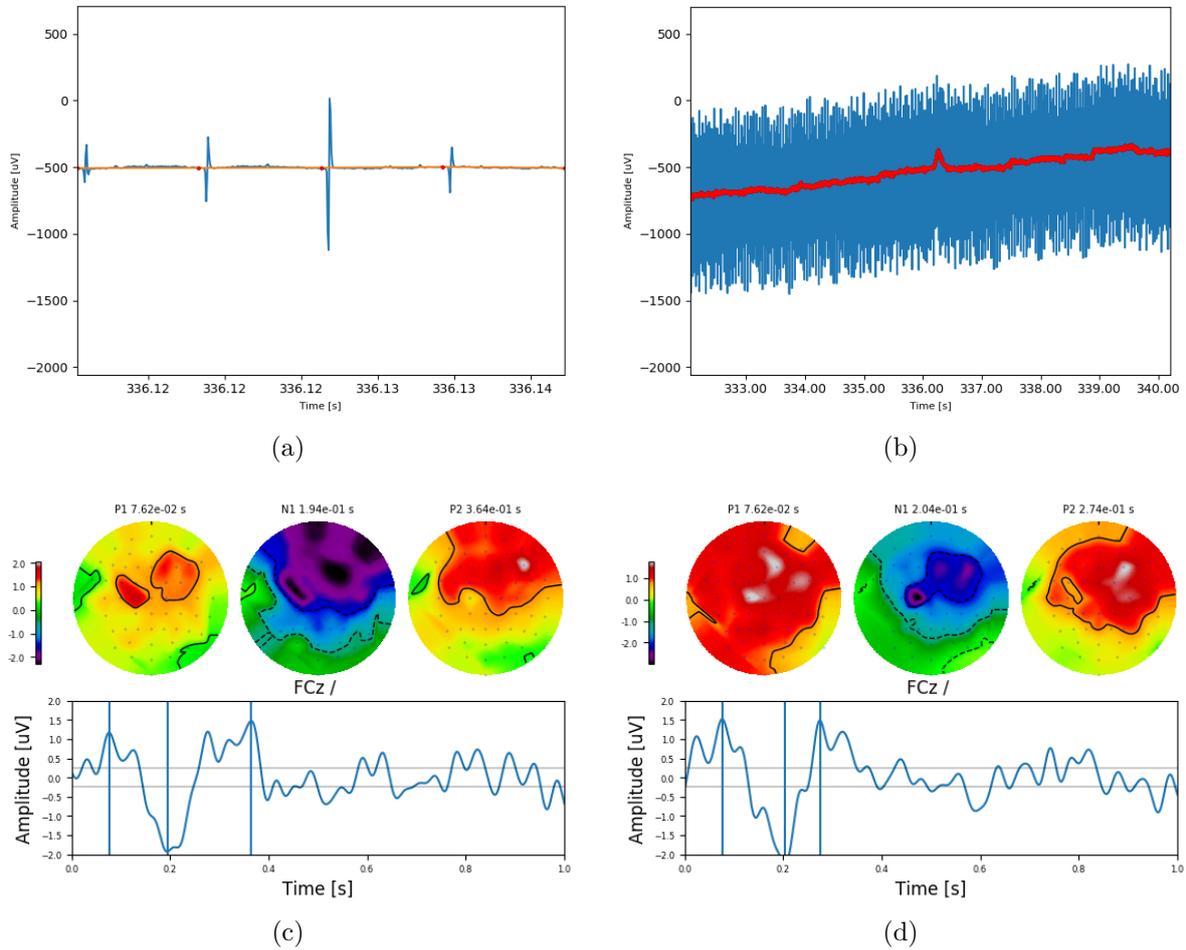


Figure 21: Top, example traces of artefact interpolation recorded from one CI participant. The original trace is shown in blue, interpolation points in red, and resulting waveform in orange. Two different time scales are shown from a) left to b) right. Bottom, example transient ACC responses to AM frequency changes imposed on a 126 pps QP-C pulse train presented on electrode 16. c), full (100%) modulated, and d) halfway (50%) modulated. The respective P1-N1-P2 peaks are shown by the vertical lines. The topographic map of each peak is shown on top where the amplitude is colour coded (in μV).

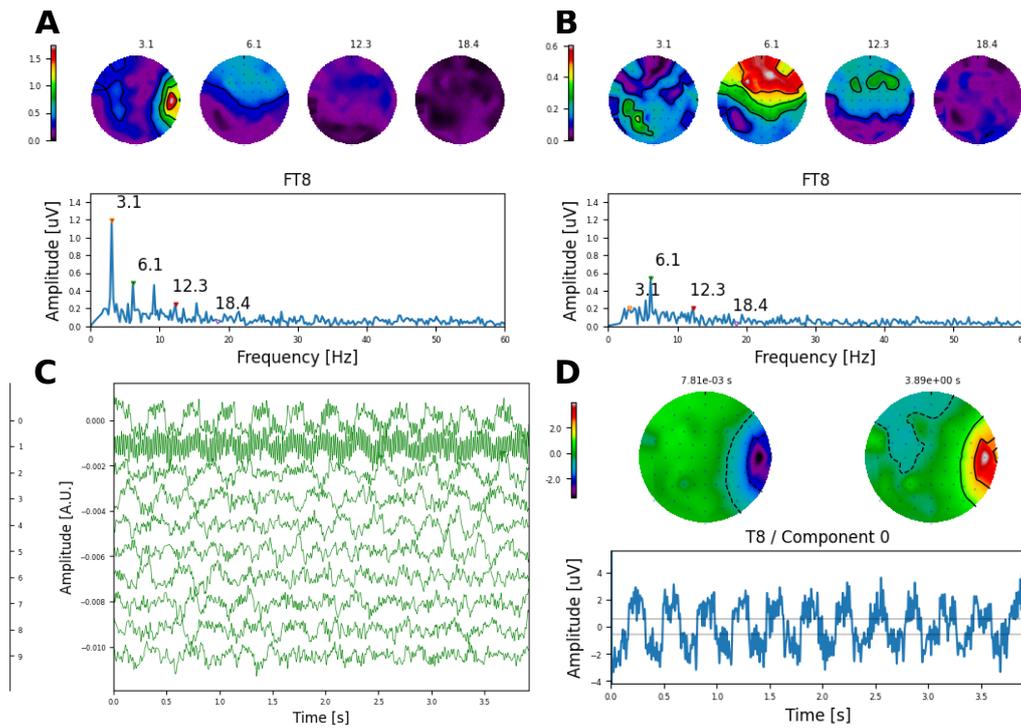


Figure 22: Example electrical artefact at a high alternating rate in a CI participant. **A** Show the original waveform and **B** the recovered response after removing the first two DSS components shown in **C**. Topographic maps show the amplitude of each frequency component indicated by the triangular markers (in μV). **D** Illustrates the projection of first DSS component on the sensor space. Topographic maps at the beginning and the end of the waveform show clearly that this component is caused by the CI (located on the right ear).