Direct cochlear recordings show human hearing nerve activity is modulated by selective attention

Gehmacher, Q.[1,2*†], Reisinger, P.[1,2†], Hartmann, T.[1,2], Keintzel, T.[3], Rösch, S.[4], Schwarz, K.[5], Weisz, N[1,2]

[1]Centre for Cognitive Neuroscience, University of Salzburg, 5020 Salzburg, Austria

[2]Department of Psychology, University of Salzburg, 5020 Salzburg, Austria

[3]Department of Otorhinolaryngology, Klinikum Wels-Grieskirchen GmbH, 4600 Wels, Austria

[4]Department of Otorhinolaryngology, Head and Neck Surgery, Paracelsus Medical University Salzburg, 5020 Salzburg, Austria

[5]MED-EL GmbH, 6020 Innsbruck, Austria

†These authors contributed equally to this work.

*Corresponding author: quirin.gehmacher@sbg.ac.at

Abstract

The architecture of the efferent auditory system enables prioritization of strongly overlapping spatiotemporal cochlear activation patterns elicited by relevant and irrelevant inputs. So far, attempts at finding such attentional modulations of cochlear activity delivered indirect insights in humans or required direct recordings in animals. The extent to which spiral ganglion cells forming the human hearing nerve are sensitive to selective attention remains largely unknown. We investigated this question by testing the effects of attending to either the auditory or visual modality on human hearing nerve activity that was directly recorded with standard commercial MED-EL cochlear implants (CI) during a stimulus-free (anticipatory) cue-target interval. When attending the upcoming auditory input, ongoing hearing nerve activity within the theta range (5-8 Hz) was enhanced. Crucially, using the broadband signal (4-25 Hz), a classifier was even able to decode the attended modality from single-trial data. Follow-up analysis showed that the effect was not driven by a narrow frequency in particular. Using direct cochlear recordings from deaf individuals, our findings suggest that cochlear spiral ganglion cells are sensitive to top-down

attentional modulations. Given the putatively broad hair-cell degeneration of these individuals, the effects are likely mediated by alternative efferent pathways as compared to previous studies using otoacoustic emissions. Successful classification of single-trial data could additionally have a significant impact on future closed-loop CI developments that incorporate real-time optimization of CI parameters based on the current mental state of the user.

Keywords: auditory, cochlear implants, hearing nerve, selective attention, top-down

Introduction

Attention describes a process by which sensory information can be prioritized. For all sensory modalities common spatiotemporal neural activity patterns have been reported on a cortical level, normally involving relevant sensory cortical regions (Frey et al., 2014; Haegens et al., 2011; Händel et al., 2010; Mazaheri et al., 2014; Salo et al., 2017; Weise et al., 2016) as well as higher level processing regions in frontal or parietal areas (Gazzaley & Nobre, 2012; Jackson et al., 2016; Nelissen et al., 2013; Weise et al., 2016; Woolgar et al., 2015). These studies suggest that selective attention is partially supported by modality-independent mechanisms. However, in contrast to e.g. the visual modality, the auditory system comprises a unique complex network of subcortical regions which all receive structural top-down connections originating in the auditory cortex (Chandrasekaran & Kraus, 2010; Elgueda & Delano, 2020; Suga, 2008; Terreros & Delano, 2015; Winer, 2006). By the most direct path, top-down information can reach the cochlea via only one extra relay through the superior olivary complex (SOC). Thus, cochlear activity can in principle be altered by top-down signals from the auditory cortex. However, studying peripheral auditory functioning requires special recording and analysis techniques (Elgueda & Delano, 2020) and has therefore been rarely investigated in this regard.

Noninvasively, evidence in humans comes from studies on otoacoustic emissions (OAEs), sounds that are generated by outer hair cell (OHC) activity in the cochlea. Activity of OHCs are modulated by a pathway from the medial olivocochlear (MOC) system that itself originates in the superior olivary complex (SOC). Spiral ganglion cells making up the auditory-nerve fibers are mainly innervated by connections of the lateral olivocochlear complex (LOC) respectively (Elgueda & Delano, 2020; Warr & Guinan, 1979). Attentional modulations of OAEs can thus be seen as a proxy for subcortical attentional modulations and have been reported for either attending the left or the right ear (Giard et al., 1994) when one out of two frequencies was task

relevant (Maison et al., 2001) or attention had to be focused on the visual or auditory modality (Dragicevic et al., 2019; Wittekindt et al., 2014). Recent work by our group (Köhler et al., in press) showed that in an audiovisual attention task cochlear activity is even modulated during a stimulus-free cue-target period, with enhanced theta-rhythmicity (~6 Hz) of spontaneous otoacoustic activity when attending to upcoming auditory input. Yet, OAEs only deliver indirect insights into the extent of the efferent auditory pathway via MOC synapses and are prone to artifacts (Francis et al., 2018). Further evidence for attention modulations of cochlear activity stems from direct recordings in animals. Delano et al. (2007) reported cochlear microphonic (CM) increases, a measure of OHC activity, together with reductions of auditory-nerve compound action potentials (CAP) in chinchillas during selective attention to visual stimuli. These effects were directly recorded with chronically implanted round-window electrodes and were later attributed to activation of MOC neurons, since electrical stimulation of MOC fibers produced opposite effects in CAP and CM amplitudes (Elgueda et al., 2011). However, whether *human* hearing nerve activity can be directly modulated via selective attention is still unknown.

While direct recordings are normally not feasible in humans, cochlear implants (CI) provide a unique opportunity for recording hearing nerve activity. CIs restore hearing by directly applying electrical stimulation to hearing nerve fibers inside the cochlea. However, the CI electrodes are used to measure hearing nerve responses to short biphasic pulses to assure hearing nerve and device functioning during and after surgery (Miller et al., 2008; Ramekers et al., 2014). These so called electrically evoked compound action potentials (ECAPs) show a first negative N1 and a second positive P1 peak (Abbas et al., 1999; He et al., 2017; Kim et al., 2010; Stypulkowski & van den Honert, 1984) within the first ~0.2 ms, corresponding to wave I of the auditory brainstem response (ABR; Christov et al., 2016), and can be measured with any conventional CI. However, the well established effects of selective attention on a cortical (Frey et al., 2014; Haegens et al., 2011; Händel et al., 2010; Mazaheri et al., 2014; Salo et al., 2017; Weise et al., 2016), as well as those reported for (corticofugal) modulations of OAEs (Dragicevic et al., 2019; Köhler et al., in press) on a cochlear level were reflected in slow oscillatory activity <30 Hz that cannot be measured with standard short-latency ECAPs. Assuming as a working hypothesis that human hearing nerve activity is modulated in a similar frequency range by selective attention, our approach appended short recording windows in a stimulus-free cue-target period. This technique allows for discrete sampling of that period within a single trial that can later be processed like standard electroencephalographic (EEG) recordings, for example, to obtain

modulations of ongoing activity in relation to attention. Interestingly, CI recipients lack the efferent MOC reflex that leads to cochlear dynamic compression in normal hearing (Lopez-Poveda et al., 2016), suggesting damaged MOC-OHC connections after substantial hair-cell degeneration. Therefore, aforementioned CI recordings during selective attention should mostly reflect modulations of spiral ganglion cell activity via LOC synaptic connections. Using an audiovisual crossmodal attention task adapted from Hartmann and Weisz (2019, see **Figure 1**), we show that ongoing hearing nerve activity in a stimulus-free cue-target interval is modulated by focused attention using standard commercial MED-EL CIs as recording devices. In addition to this average condition-level effect, we show that a classifier is even able to decode attended modality on a single-trial basis, which could additionally have important implications for the use of conventional CIs in a closed-loop system.

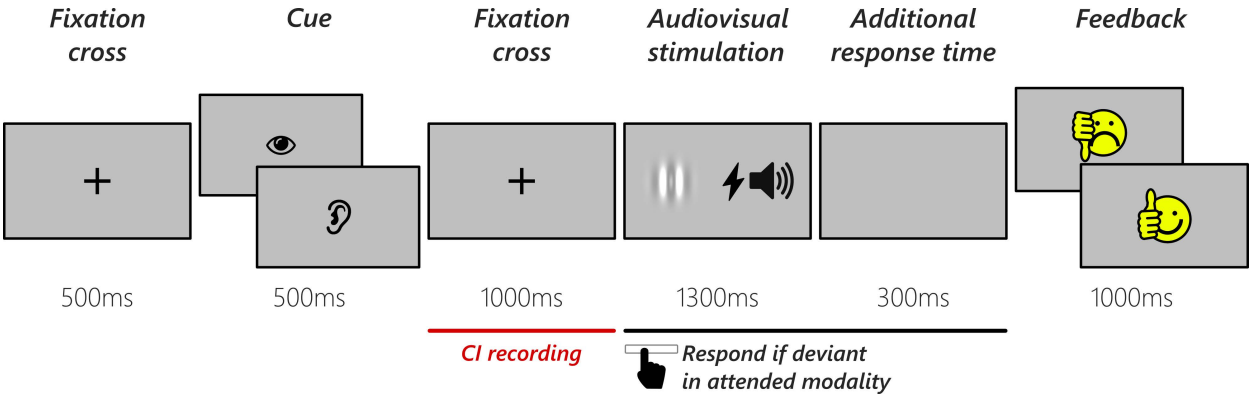

Figure 1. *Schematic illustration of the crossmodal attention task.* Each trial started with a fixation cross, followed by a cue indicating either to attend the visual or auditory domain. A second fixation cross appeared and an auditory and visual stimulus were presented afterwards. When the stimulus in the attended modality was deviant (visual: gabor patch tilt, auditory: oddball sound), participants had to respond by pressing the spacebar. The additional response time accounted for trials where the gabor patch tilted towards the end of the stimulation. At the end of each trial, feedback was given in the form of a smiley face. The red line denotes the time window where hearing nerve activity was recorded via the CI.

Results

Sixteen CI users performed a crossmodal attention task (similar to Hartmann & Weisz, 2019) where attention had to be focused on an upcoming auditory or visual stimulus (see **Figure 1**). Hearing nerve activity was recorded directly via one of their standard CI electrodes in the stimulus-free cue-target interval. We calculated the power spectral density of the signal and

compared the two conditions (attend auditory vs. visual) in the theta and alpha band. Afterwards, a classifier was utilized to decode the attended modality on a single-trial basis using the broadband signal and frequency bands typically associated with selective attention (theta, alpha, beta).

Behavioral results

Participants gave a correct response in 96% (*SD* = 2.7%) of all trials. The number of correct trials did not differ significantly between the two conditions, according to a dependent sample t-test (auditory: *M* = 245 (*SD* = 9.8); visual: *M* = 242 (*SD* = 8.8); $t(15) = 1.32$, $p = 0.21$, $d = 0.33$). When there was an oddball in the cued domain, a correct response was given in 75% (*SD* = 19.0%) of the trials. In the auditory condition, the percentage of correct oddball trials was 72% (*SD* = 30.8%) and in the visual condition 78% (*SD* = 14.2%). Overall, the behavioral findings suggested that participants performed the task in a compliant manner.

Human hearing nerve activity is modulated by selective attention

In a first analysis step, we calculated the broadband PSD from 4-25 Hz, separately for each condition (attend auditory/attend visual). The resulting power spectra (see **Figure 2A**) by themselves showed no clear peaks, however the grand average condition contrast spectrum indicates differences that are mainly centered in distinct frequency ranges. Based on previous OAE and M/EEG work (Köhler et al., in press; Mazaheri et al., 2014), we statistically compared the two conditions in the theta (averaged between 5-8 Hz) and alpha frequency band (averaged between 9-13 Hz; **Figure 2B**). A cluster-based permutation test in the theta frequency band showed that prestimulus power is higher when attending the auditory domain ($p = 10.00e^{-05}$, $d = 0.49$). No cluster was found in the alpha frequency band ($p = 1.00$, $d = 0.20$). Given the distribution of the individual average power in both FOI, a rather high interindividual variability can be seen.

The results so far showed that selective attention modulates directly recorded cochlear activity, with the effect being in particular pronounced in the theta frequency range: attending to an upcoming auditory stimulus resulted in higher power recorded from the CI electrode.

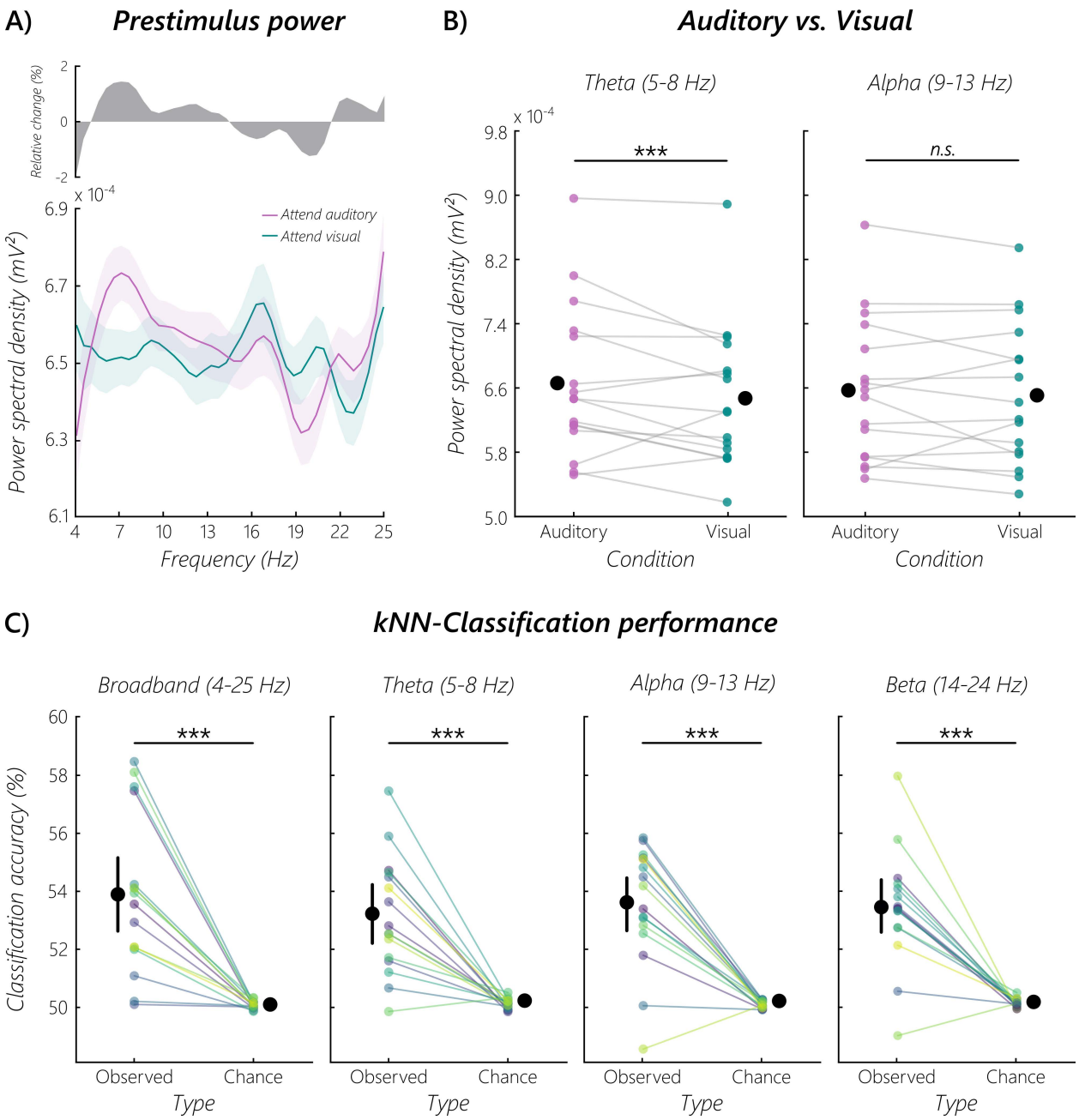

Figure 2. *Prestimulus power modulations and decoding of selective attention.* (A) Grand average prestimulus power spectra from 4-25 Hz when attending the auditory or visual domain. The top panel indicates the relative change between the auditory and visual domain. The shaded areas in the bottom panel represent the standard error of the mean for within-subjects designs (O'Brien & Cousineau, 2014). (B) Average prestimulus power in the theta and alpha band, separated by the two conditions. Black dots indicate the group mean for the respective condition. A cluster-based permutation test in the averaged theta FOI resulted in a statistically significant difference when testing the hypothesis that performance is higher when attending the auditory domain (p = 10.00e^{-05}, d = 0.49). No cluster was found in the alpha band (p = 1.00, d = 0.20). The asterisks indicate a statistically significant difference (n.s. = not significant). (C) A kNN-Classifier was used to decode attended modality from single-trial prestimulus power spectra. Resulting *Observed* accuracies were contrasted with respective *Chance* levels of a random permutation

test for all FOI. Contrasts revealed significant ($p < 0.001$) decoding performance throughout spectra with fairly similar effects (broadband: $t(15) = 5.60$, $p = 2.60e^{-05}$, $d = 1.96$; theta: $t(15) = 5.83$, $p = 1.70e^{-05}$, $d = 2.11$; alpha: $t(15) = 6.78$, $p = 3.00e^{-06}$, $d = 2.34$; beta: $t(15) = 6.40$, $p = 6.00e^{-06}$, $d = 2.33$) on a group level (represented by black dots; error bars = 95% CI). However, on a single-subject level the attention effect was most pronounced for individually specific FOI, resulting in significant above chance decoding for 12 out of 16 subjects. Every subject is represented by the same color in all four FOI columns.

Attended modality can be decoded from single-trial CI recordings

We used prestimulus power spectra for a kNN-Classifier to show that attention modulation of ongoing hearing nerve activity in humans is even reflected in single-trial CI recordings. To ensure that the classifier was able to differentiate hearing nerve activity when attending the auditory compared to the visual domain in general, we calculated a t-test between *Observed* classification accuracies and respective *Chance* levels of broadband power spectra, showing that this attention effect was decodable significantly above chance ($t(15) = 5.60$, $p = 2.60e^{-05}$, $d = 1.96$; **Figure 2C**). Given the significant difference over a broad frequency range, we were further interested in whether this attention effect was driven by one of the FOI usually connected with selective attention in OAE and M/EEG studies. We therefore calculated a two-factor repeated measures ANOVA to compare the effect of selective attention on kNN-Classification accuracy for different FOI (broadband, theta, alpha, beta) and Types (*Observed*, *Chance*). Results show no significant effect of FOI ($F(3, 45) = 0.37$, $p = 0.78$, $\eta_p^2 = 0.02$), yet show a significant effect for *Observed* vs. *Chance* accuracies ($F(1, 15) = 136.55$, $p = 6.21e^{-09}$, $\eta_p^2 = 0.90$), with higher accuracies for *Observed* ($M = 0.53$) than *Chance* ($M = 0.50$) levels. No FOI x Type interaction on decoding results was found ($F(3, 45) = 0.36$, $p = 0.78$, $\eta_p^2 = 0.02$). As a main effect of Type and no interaction of FOI x Type indicated that selective attention can be decoded from all FOI separately, in addition we computed three t-tests for theta, alpha, and beta bands contrasting respective *Observed* and *Chance* levels. For all three FOI a significant difference was found for selective attention decoding (theta: $t(15) = 5.83$, $p = 1.70e^{-05}$, $d = 2.11$; alpha: $t(15) = 6.78$, $p = 3.00e^{-06}$, $d = 2.34$; beta: $t(15) = 6.40$, $p = 6.00e^{-06}$, $d = 2.33$; **Figure 2C**). Additionally, random permutation tests of kNN classification within all four different FOI gave insights into single-subject decoding performance across the different frequency spectra. Independent of FOI, an overall number of 12 subjects (i.e. 75% of the sample) showed significant ($p < 0.05$) above chance decoding of focused attention during the stimulus-free cue-target interval.

Discussion

The efferent auditory system comprises a complex arrangement of subcortical and peripheral pathways where - uniquely among the senses - the cochlear activity can be altered by top-down signals from the auditory cortex via the SOC (Chandrasekaran & Kraus, 2010; Elgueda & Delano, 2020; Suga, 2008; Terreros & Delano, 2015; Winer, 2006). Profound evidence supports the notion of altered oscillatory neural activity by selective attention on a cortical level (Frey et al., 2014; Haegens et al., 2011; Händel et al., 2010; Mazaheri et al., 2014; Salo et al., 2017; Weise et al., 2016), prioritizing attended over unattended events in the environment which is predominantly reflected in modulations within the alpha band (~9-13 Hz; Frey et al., 2014; Sauseng et al., 2005; for a review see Foxe & Snyder, 2011; notably also Antonov et al., 2020 that challenges current views). While such changes have also been reported for the auditory cortex (Weisz et al., 2011), much less is known for subcortical structures along the efferent pathway. This is in particular the case when it comes to the *human* cochlea as special recording and analysis techniques are required (Elgueda & Delano, 2020). So far, investigating attentional modulation of cochlear activity in humans had to rely on indirect recordings of OAEs; a noninvasive approach for measuring OHC activity that receives modulatory signals from the SOC through its medial neurons (MOC). Recent evidence suggests slow modulations of cochlear activity when attention had to be focused on the auditory modality. Dragicevic et al. (2019) were the first ones to establish attention-related modulations of distortion product OAE (DPOAE) activity in a low (<10 Hz) frequency range. Going beyond these stimulus-evoked changes, even when no auditory stimulus is presented, cochlear activity is enhanced in an analogous frequency range as shown by Köhler et al. (in press). These striking results provide substantial evidence for a notable role of selective attention at the very first stages of sound processing in the human auditory system. However, in addition to an indirect measurement approach that is prone to artifacts (Francis et al., 2018), studying OAEs cannot address direct modulation of hearing nerve activity since spiral ganglion cells are efferently innervated by a separate pathway. To our knowledge, this attentional modulation via the LOC remains completely unknown as direct recordings of hearing nerve activity are normally not feasible in humans. Given the damage to OHCs and their connections to the MOC in CI recipients (Lopez-Poveda et al., 2016) potential alterations of respective hearing nerve activity in a selective attention paradigm should largely reflect top-down signals from the LOC. Our results, contrasting direct cochlear recordings during auditory and visual attention, show that ongoing

hearing nerve activity is top-down modulated, putatively suggesting a role of the LOC pathway in selective attention.

While cortical and OAE-based measures suggest attention-related effects in distinct frequency bands (Köhler et al., in press; Mazaheri et al., 2014), our results are mixed in this respect. The broadband frequency analysis of the prestimulus interval showed no clear peaks (**Figure 2A**). This, however, may also be the result of low signal-to-noise ratios (SNRs), as commercial CIs are so far not optimized to do these kinds of continuous electrophysiological recordings. Indeed a grand average of the condition differences points to maximal effects in a frequency range overlapping with the one reported by Köhler et al. (in press). When analyzing the prestimulus power spectra in two FOI (theta and alpha), we found a selective attention effect in theta resulting in enhanced power while attending to the auditory modality (**Figure 2B**). This result corroborates our previous finding using a similar paradigm, where ongoing OAEs in the theta band (~6 Hz) were enhanced while attending an upcoming auditory stimulus (Köhler et al., in press). We found no selective attention effect in the alpha band in concordance with aforementioned studies of otoacoustic activity. It is therefore possible that this frequency band does not play a central role in selective attention at the peripheral level. Further studies with an optimized recording setup will be necessary to address this issue.

Building upon conventional analyses of condition-level fast Fourier transform (FFT) averages, we decided to use single-trial frequency spectra to classify anticipatory attentional focus during the stimulus-free cue-target period. With this approach we aimed to get more detailed insight into fine-grained differences between attentional states coded within modulations of direct cochlear recordings that could be missed by condition-level averaging approaches and indirect OAE measurements. Strikingly, classification of the broadband signal (4-25 Hz) revealed significantly improved differentiation of attended modality compared to the average condition-level effect of the FFT results (**Figure 2C**). There seems to be sufficient top-down modality specific information entailed in single-channel CI recordings to ensure above chance classification performance without even specifying a particular attention-related frequency band - a clear improvement to the standard FFT approach. As a matter of fact, follow-up analysis showed that the performance was not driven by one of three FOI (theta, alpha, beta) usually associated with selective attention, but instead it revealed that the contribution of each of these frequency bands to broadband classification was fairly similar. However, the decoding approach allowed for additional insight into single-subject classification performance and showed high

interindividual variability in terms of an optimal spectral frequency band. It remains to be determined whether this effect is driven by local idiosyncrasies at the peripheral level (e.g. synaptic connections between LOC and spiral ganglion cells) or even involves particular activity patterns at higher hierarchical levels. Independent of the precise origins of our effects observed at the hearing nerve, the decoding results open up avenues to future developments towards closed-loop CIs that incorporate mental states of the recipient reflected in cochlear activity into adaptive stimulation in real time. Attempts so far with standard modern CIs required exhaustive recording and analysis sessions and indicated additional hard- and software modifications (Mc Laughlin et al., 2012) in order to extract meaningful neuronal activity with additional extracochlear (i.e. EEG) electrodes, especially for real time applications. In the present study, we used standard commercial MED-EL CIs to record ongoing activity directly at the hearing nerve during a stimulus-free anticipatory cue-target interval without additional electrodes or implantations. As we show, a classifier could use the frequency information of this signal to anticipate the attentional state of the recipient. Future research will need to address which cognitive states can be decoded directly at the hearing nerve and how this information could be exploited in a closed-loop CI setup.

This study shows that individuals with a CI form a model population to deepen our understanding of how cognition can lever the efferent auditory system to modulate auditory input at the earliest stages of processing. In the present study we focused on stimulation-free anticipatory cue-target periods using a popular experimental setup in cognitive neuroscience (Hartmann & Weisz, 2019; Köhler et al., in press; Posner, 1980; Wittekindt et al., 2014). The extent to which the identified top-down modulations also play a role in more complex environments with natural speech, for example in attending one of multiple speakers at a “cocktail-party”, should be addressed in further studies and could add valuable information to an ongoing debate on speech processing in such scenarios. As CI recipients show highly varying listening success in multitalker situations (Loizou et al., 2009), their ability to (re-)engage the path from the LOC to auditory-nerve fibers in selective processing could play a key role in coping post-operatively with those situations. Future studies could combine our direct cochlear recording approach with EEG-based measurements of cortical modulations during selective attention. However, so far technical restrictions by the device itself (Abbas et al., 2017; Campbell et al., 2015; Tejani et al., 2019) and superimposing artifacts that would be caused by the stimulation with complex sounds like speech don’t allow researchers to use CI electrodes for

simultaneous stimulation and recording of hearing nerve activity and needed to be optimized for such continuous applications in future developments. Nonetheless, our approach shows that ongoing hearing nerve activity contains relevant and, crucially, classifiable information about current attentional status that could be used to adapt the CI's processing strategy, resulting in less fatigue and higher intelligibility during challenging listening situations.

In summary, to the best of our knowledge our study is the first to investigate attentional effects on activity recorded directly from the hearing nerve in humans. We confirm and extend previous indirect measurements, suggesting attentional modulations in the theta frequency range. Importantly, we also show that selective attention can be decoded above chance at a single-trial and even individual level. Previous reports on attentional modulations of cochlear activity relied on OAEs, which are driven by the MOC pathway. Our results strongly suggest that the LOC pathway can also be exploited in a top-down down fashion to affect spiral ganglion cells directly.

Materials and Methods

Participants

21 right-handed CI users (4 females, M_{age} = 57.5, SD_{age} = 11.9) participated in the study, all with a minimum CI experience of six months. Participants were recruited via the ENT departments of the hospitals in Salzburg (n = 10) and Wels-Grieskirchen (n = 11). Three participants were excluded because of a too weak contact between transmitting CI coil and receiver that was required for the study. One participant showed no N1 in recorded ECAPs, which could indicate a measurement problem and was therefore excluded. One participant quit during the session due to concentration problems. This led to a final sample size of 16 participants (3 females, M_{age} = 53.8, SD_{age} = 12.0; see Table S1 for further details). All participants reported no previous neurological or psychiatric disorders, and reported normal or corrected-to-normal vision. All participants signed an informed consent and were reimbursed with 10 Euro per hour. The experimental protocol was approved by the ethics committee of the University of Salzburg and was carried out in accordance with the Declaration of Helsinki.

Stimuli and Procedure

The experimental procedure was implemented in MATLAB 8.6 (The MathWorks Inc., Natick, Massachusetts, USA) using custom scripts. Presentation of visual stimuli and response

collection was achieved with a previous version (th_ptb; https://gitlab.com/thht/th_ptb) of the Objective Psychophysics Toolbox (o_ptb; Hartmann & Weisz, 2020), which adds an additional class-based abstraction layer in addition to the Psychophysics Toolbox (Version 3.0.14; Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Cochlear stimulation as well as recording was performed via the MAX Programming Interface, a device which is part of the clinical standard setup that enables control of the implant, together with the Research Interface Box 2 Dynamic-link library (RIB2 DLL provided by the University of Innsbruck, Innsbruck, AT; Litovsky et al., 2017). To ensure accurate stimulus presentation and triggering, timings were measured with the Black Box ToolKit v2 (The Black Box ToolKit Ltd., Sheffield, UK).

Participants were seated in front of a computer screen and were asked to remove their CI processor and coil to replace it with a coil connected to the MAX Programming Interface. For bilateral CI users, the side with the better subjective hearing performance and/or longer implantation date was used (see Table S1). Primarily the CI coil model MAX Coil was used, but if the magnet was too weak to ensure a stable connection, the CI coil model MAX Coil S was used. As a first step, the individual electrical hearing threshold was determined with a standard tone with a stimulation frequency of 100 Hz and a duration of 300 ms. To ensure that the auditory stimulation was at a comfortable level during the experiment, the individual maximum loudness was determined, for the standard and an oddball tone respectively. An oddball tone with the maximum possible stimulation frequency of 9990 Hz (based on the used phase duration of 30 µs per phase for sequential biphasic pulses) and a duration of 300 ms was used. The described routines were implemented using custom scripts and the Palamedes Toolbox (Prins & Kingdom, 2018). Afterwards, as a functionality check of the measurement setup, ECAPs were recorded (for further details, see Methods S1; Bahmer et al., 2010). For the crossmodal attentional task described later, it was necessary that two stimulation frequencies could be distinguished. Because of interindividual differences when hearing with a CI, it was necessary to adjust these stimulation frequencies for every participant. Participants were asked, after hearing a standard and oddball tone, if the first or the second tone had a higher stimulation frequency. The standard tone had a stimulation frequency of 100 Hz and a duration of 300 ms. The initial stimulation frequency of the oddball tone (also with a duration of 300 ms) was determined by the results of the aforementioned maximum loudness procedure. This procedure was carried out using a Bayesian active sampling protocol to estimate the model parameters of the psychometric function (Kontsevich & Tyler, 1999; Sanchez et al., 2016) and was

implemented with the VBA Toolbox (Daunizeau et al., 2014). To define the individual oddball stimulation frequency for the subsequent crossmodal attention task, the algorithm searched for the optimal difference in logarithmic steps from 1 to 9890 Hz and this value was subsequently added to the standard stimulation frequency. Six participants heard no clear difference and it was necessary to adjust the oddball stimulation frequency manually, with values between 114 and 600 Hz.

The actual experiment was carried out as a crossmodal attention task (see **Figure 1**; similar to Hartmann & Weisz, 2019) in six blocks, with 85 trials per block. Each trial started with a 500 ms fixation cross, followed by a cue that indicated either to attend the auditory or the visual modality. Every block had 43 auditory and 42 visual cues. The cue was a picture of an eye or ear, presented for 500 ms. A second fixation cross appeared for 1000 ms and the audiovisual stimulation started afterwards. The auditory stimulation consisted of a 300 ms tone with a stimulation frequency of 100 Hz and was directly presented via the CI coil. The visual stimulation was a vertically oriented gabor patch (spatial frequency: 0.01 cycles/pixel, sigma: 60 pixels, phase: 90°), presented for 1300 ms in the center of the screen. In every block, 8 trials were randomly chosen as visual oddball trials. Independently, another 8 trials were chosen to be auditory oddball trials. Therefore it was possible that a trial was a visual and auditory oddball trial simultaneously. In visual oddball trials, the gabor patch tilted 10° to the left, with a random onset. In auditory oddball trials, a 300 ms tone with the individual oddball stimulation frequency was presented. Participants had to press the spacebar if the current trial had an oddball in the cued domain. To account for trials where the visual oddball onset was towards the end of the stimulation, an additional response time of 300 ms was provided. After each trial, feedback in the form of a smiley face displayed for 1000 ms, indicated if the response was correct or not. To ensure correct understanding of and response during the task, participants completed one block as a practice run before the actual experiment. The total duration of the experiment was about 90 minutes including breaks and preparation.

Recording of hearing nerve activity

We exploit the ability of CIs to record electrical activity from the cochlea in short time windows, but in contrast to previous approaches (Abbas et al., 2017; Mc Laughlin et al., 2012), in a stimulus-free cue-target period. Using a custom developed MATLAB toolbox to abstract MAX Programming Interface commands, we recorded hearing nerve activity via the CI electrode (for

bilateral CI users, see Table S1 which side was used). In every participant, the first (i.e. most apical) electrode was used for the recordings. Each recording window was 1.7 ms long, followed by a 13.68 ms reset period resulting in a sampling frequency of 65 Hz (1.7 ms recording + 13.68 ms reset time). The technical specifics of the measurement system added a random offset to each of the recordings (Gaussian noise, *SD* = 0.4 mV). Because of the USB connection between the computer and the MAX Programming Interface, the start of the first recording window had a jitter of 27 ms, but the system sent a highly precise trigger when it started. Due to technical limitations, it was not possible to record and stimulate simultaneously. We performed recordings in the 1000 ms prestimulus window (see red line in **Figure 1**).

Data preprocessing

The raw data was analyzed in MATLAB 9.8 (The MathWorks, Natick, Massachusetts, USA). Due to filter artifacts (using the standard filter from the used RIB2 package), the first 100 samples (= 0.083 ms) from every recording window were discarded. Afterwards, the recording was averaged and treated as one sample point. By repeating these steps for every window and concatenating the single samples, a recording length of 1 second with a sampling frequency of 65 Hz was reached. The data was further preprocessed with the FieldTrip toolbox (revision ea6897bb7; Oostenveld et al., 2010) and a bandpass filter between 4 and 25 Hz was applied (hamming-windowed sinc FIR filter, onepass-zerophase, order: 424, transition width: 0.5 Hz). For one participant, 15 trials had to be rejected because the CI coil fell off during the last trials of one block. Only trials with a correct response were analyzed, which were on average 488 trials (*SD* = 15.8). The number of correct trials was not significantly different between the two conditions (see Behavioral results).

Frequency analysis

Next, data was demeaned, detrended and power spectral density (PSD) from 4 to 25 Hz was computed on the whole 1000 ms prestimulus window ('mtmfft' implementation in FieldTrip with a Hann window) separately for the two conditions. For **Figure 2A**, no bandpass filter was applied, condition-specific power spectra were smoothed (five-point moving average), grand-averaged, and corrected error bars for within-subjects designs were calculated (O'Brien & Cousineau, 2014).

We defined two frequency bands of interest (FOI), theta (5-8 Hz) and alpha (9-13 Hz). Theta was selected because of previous work on OAEs that showed attentional modulations in this FOI (Dragicevic et al., 2019; Köhler et al., in press). On a cortical level, previous work showed that auditory alpha activity reflects attentional processes (Frey et al., 2014; Mazaheri et al., 2014; Müller & Weisz, 2012; Weise et al., 2016; Weisz et al., 2014; for a review see Weisz et al., 2011). Therefore, we decided to analyze this FOI at the cochlear level.

To test the hypothesis that power was higher when the auditory domain was attended, statistical testing of PSD was performed with a cluster-based permutation test (dependent samples t-test, 10000 randomizations, one-tailed; Maris & Oostenveld, 2007). We averaged the theta and alpha FOI and tested them separately.

Decoding analysis

For decoding of attended modality on a single-trial basis, we performed k-nearest neighbors (kNN) classification of single-trial power spectra using scikit-learn (Version 0.23.1 running on Python 3.7.7; Pedregosa et al., 2011) separately for a broadband signal (4-25 Hz) followed by standard frequency bands associated with selective attention (theta: 5-8 Hz, alpha: 9-13 Hz, beta: 14-24 Hz). We decided to use the kNN classification approach as data was recorded from a single CI channel over a one second period resulting in low numbers of features (i.e. frequency points per band), a classification problem usually solved better by a kNN approach (Eisa et al., 2018). At first, a subject's data was standardized to unit variance and zero mean. For the classification process of each subject, the best number of neighbors was determined by searching the hyper-parameter space for the best cross-validation (CV) score of a kNN model using the implemented *GridSearchCV* function with a 2-fold CV on shuffled class samples (*StratifiedKFold(shuffle=True)*) that was fit to the data for every FOI. Our decision for a 2-fold CV was based on recommendations in case of low sample / effect size data (Jamalabadi et al., 2016). The numbers of neighbors to use during the gridsearch were defined as ranging from one to 10% of trials in the dataset in odd numbers (1, trials/10, stepsize=2) to avoid the conflict of even neighbors in a two-class problem (attend auditory vs. visual). Given the novel approach, we could not exclude that the classifier would pick up on a few outlying data points. In order to address this issue explicitly, the classifier was tested on the same noisy data, albeit with randomly shuffled condition labels. Samples were thus classified and tested for significance with

the best scoring number of neighbors in a 1000 random permutation test and the 2-fold CV procedure.

The resulting *Observed* and *Chance* accuracy values (where chance level was calculated as the mean accuracy of the 1000 random permutation scores) for every FOI were then statistically tested using pingouin (Version 0.3.8 running on Python 3.7.7; Vallat, 2018). In a first step, to test whether hearing nerve modulation was generally reflected within classification results, broadband values (*Observed* vs. *Chance*) were compared using a one-sided t-test. Then, classification results of all four FOI were compared in a two-factor repeated measures ANOVA with the factors *FOI* (broadband, theta, alpha, beta) and *Type* (Observed vs. Chance) to check whether the attention effect was driven by one of the predefined FOI. Finally, theta, alpha, and beta bands were also tested for significant differences during focused attention computing three one-sided t-tests with respective values (*Observed* vs. *Chance*).

Acknowledgments

The authors would like to thank Agnes Koller and Lisa Niederwanger for their help in recruiting participants. Quirin Gehmacher and Patrick Reisinger are supported by the Austrian Research Promotion Agency (FFG; BRIDGE 1 project “SmartCIs”; 871232) and the Austrian Science Fund (FWF; Doctoral College "Imaging the Mind"; W 1233-B).

Author Contributions

Q.G. and P.R. collected and analyzed data, generated the figures, and wrote the manuscript. T.H. designed the experiment, collected and analyzed data, developed analysis methods, and edited the manuscript. T.K. and S.R. recruited participants and consulted on the clinical aspects. K.S. developed analysis methods, provided input on data analysis, and edited the manuscript. N.W. acquired the funding, supervised the project, and edited the manuscript.

Declaration of Interests

K.S. is an employee of MED-EL GmbH. All other authors declare no competing interests.

References

Abbas, P. J., Brown, C. J., Shallop, J. K., Firszt, J. B., Hughes, M. L., Hong, S. H., & Staller, S. J. (1999). Summary of Results Using the Nucleus CI24M Implant to Record the Electrically Evoked Compound Action Potential. *Ear and Hearing*, *20*(1), 45–59. https://doi.org/10.1097/00003446-199902000-00005

Abbas, P. J., Tejani, V. D., Scheperle, R. A., & Brown, C. J. (2017). Using Neural Response Telemetry to Monitor Physiological Responses to Acoustic Stimulation in Hybrid Cochlear Implant Users. *Ear and Hearing*, *38*(4), 409. https://doi.org/10.1097/AUD.0000000000000400

Antonov, P. A., Chakravarthi, R., & Andersen, S. K. (2020). Too little, too late, and in the wrong place: Alpha band activity does not reflect an active mechanism of selective attention. *NeuroImage*, *219*, 117006. https://doi.org/10.1016/j.neuroimage.2020.117006

Bahmer, A., Peter, O., & Baumann, U. (2010). Recording and analysis of electrically evoked compound action potentials (ECAPs) with MED-EL cochlear implants and different artifact reduction strategies in Matlab. *Journal of Neuroscience Methods*, *191*(1), 66–74. https://doi.org/10.1016/j.jneumeth.2010.06.008

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Campbell, L., Kaicer, A., Briggs, R., & O'Leary, S. (2015). Cochlear Response Telemetry: Intracochlear Electrocochleography via Cochlear Implant Neural Response Telemetry Pilot Study Results. *Otology & Neurotology*, *36*(3), 399. https://doi.org/10.1097/MAO.0000000000000678

Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*(2), 236–246.

https://doi.org/10.1111/j.1469-8986.2009.00928.x

Christov, F., Munder, P., Berg, L., Bagus, H., Lang, S., & Arweiler-Harbeck, D. (2016). ECAP analysis in cochlear implant patients as a function of patient's age and electrode-design. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *133*, S1–S3. https://doi.org/10.1016/j.anorl.2016.04.015

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLOS Computational Biology*, *10*(1), e1003441. https://doi.org/10.1371/journal.pcbi.1003441

Delano, P. H., Elgueda, D., Hamame, C. M., & Robles, L. (2007). Selective Attention to Visual Stimuli Reduces Cochlear Sensitivity in Chinchillas. *Journal of Neuroscience*, *27*(15), 4146–4153. https://doi.org/10.1523/JNEUROSCI.3702-06.2007

Dragicevic, C. D., Marcenaro, B., Navarrete, M., Robles, L., & Delano, P. H. (2019). Oscillatory infrasonic modulation of the cochlear amplifier by selective attention. *PLOS ONE*, *14*(1), e0208939. https://doi.org/10.1371/journal.pone.0208939

Eisa, D. A., Taloba, A. I., & Ismail, S. S. I. (2018). A Comparative Study on using Principle Component Analysis with different Text Classifiers. *International Journal of Computer Applications*, *180*(31), 1–6. https://doi.org/10.5120/ijca2018916800

Elgueda, D., & Delano, P. H. (2020). Corticofugal modulation of audition. *Current Opinion in Physiology*, *18*, 73–78. https://doi.org/10.1016/j.cophys.2020.08.016

Elgueda, D., Delano, P. H., & Robles, L. (2011). Effects of Electrical Stimulation of Olivocochlear Fibers in Cochlear Potentials in the Chinchilla. *Journal of the Association for Research in Otolaryngology*, *12*(3), 317–327. https://doi.org/10.1007/s10162-011-0260-9

Foxe, J. J., & Snyder, A. C. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, *2*.

https://doi.org/10.3389/fpsyg.2011.00154

Francis, N. A., Zhao, W., & Guinan Jr., J. J. (2018). Auditory Attention Reduced Ear-Canal Noise in Humans by Reducing Subject Motion, Not by Medial Olivocochlear Efferent Inhibition: Implications for Measuring Otoacoustic Emissions During a Behavioral Task. *Frontiers in Systems Neuroscience*, *12*. https://doi.org/10.3389/fnsys.2018.00042

Frey, J. N., Mainy, N., Lachaux, J.-P., Müller, N., Bertrand, O., & Weisz, N. (2014). Selective Modulation of Auditory Cortical Alpha Activity in an Audiovisual Spatial Attention Task. *Journal of Neuroscience*, *34*(19), 6634–6639. https://doi.org/10.1523/JNEUROSCI.4813-13.2014

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Giard, M.-H., Collet, L., Bouchet, P., & Pernier, J. (1994). Auditory selective attention in the human cochlea. *Brain Research*, *633*(1), 353–356. https://doi.org/10.1016/0006-8993(94)91561-X

Haegens, S., Nácher, V., Luna, R., Romo, R., & Jensen, O. (2011). α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences*, *108*(48), 19377–19382. https://doi.org/10.1073/pnas.1117190108

Händel, B. F., Haarmeier, T., & Jensen, O. (2010). Alpha Oscillations Correlate with the Successful Inhibition of Unattended Stimuli. *Journal of Cognitive Neuroscience*, *23*(9), 2494–2502. https://doi.org/10.1162/jocn.2010.21557

Hartmann, T., & Weisz, N. (2019). Auditory cortical generators of the Frequency Following Response are modulated by intermodal attention. *NeuroImage*, *203*, 116185.

https://doi.org/10.1016/j.neuroimage.2019.116185

Hartmann, T., & Weisz, N. (2020). An introduction to the Objective Psychophysics Toolbox. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.585437

He, S., Teagle, H. F. B., & Buchman, C. A. (2017). The electrically evoked compound action potential: From laboratory to clinic. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00339

Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2016). Feature-selective Attention in Frontoparietal Cortex: Multivoxel Codes Adjust to Prioritize Task-relevant Information. *Journal of Cognitive Neuroscience*, *29*(2), 310–321. https://doi.org/10.1162/jocn_a_01039

Jamalabadi, H., Alizadeh, S., Schönauer, M., Leibold, C., & Gais, S. (2016). Classification based hypothesis testing in neuroscience: Below-chance level classification rates and overlooked statistical properties of linear parametric classifiers. *Human Brain Mapping*, *37*(5), 1842–1855. https://doi.org/10.1002/hbm.23140

Kim, J.-R., Abbas, P. J., Brown, C. J., Etler, C. P., O'Brien, S., & Kim, L.-S. (2010). The Relationship Between Electrically Evoked Compound Action Potential and Speech Perception: A Study in Cochlear Implant Users With Short Electrode Array. *Otology & Neurotology*, *31*(7), 1041–1048. https://doi.org/10.1097/MAO.0b013e3181ec1d92

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 1–16.

Köhler, M. H. A., Demarchi, G., & Weisz, N. (in press). Cochlear activity in silent cue-target intervals shows a theta-rhythmic pattern and is correlated to attentional alpha and theta modulations. *BMC Biology*.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope

and threshold. *Vision Research*, *39*(16), 2729–2737. https://doi.org/10.1016/S0042-6989(98)00285-5

Litovsky, R. Y., Goupell, M. J., Kan, A., & Landsberger, D. M. (2017). Use of Research Interfaces for Psychophysical Studies With Cochlear-Implant Users. *Trends in Hearing*, *21*, 2331216517736464. https://doi.org/10.1177/2331216517736464

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., & Roland, P. (2009). Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *The Journal of the Acoustical Society of America*, *125*(1), 372–383. https://doi.org/10.1121/1.3036175

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., & Wilson, B. S. (2016). A Binaural Cochlear Implant Sound Coding Strategy Inspired by the Contralateral Medial Olivocochlear Reflex. *Ear and Hearing*, *37*(3), e138. https://doi.org/10.1097/AUD.0000000000000273

Maison, S., Micheyl, C., & Collet, L. (2001). Influence of focused auditory attention on cochlear activity in humans. *Psychophysiology*, *38*(1), 35–40. https://doi.org/10.1111/1469-8986.3810035

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mazaheri, A., van Schouwenburg, M. R., Dimitrijevic, A., Denys, D., Cools, R., & Jensen, O. (2014). Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *NeuroImage*, *87*, 356–362. https://doi.org/10.1016/j.neuroimage.2013.10.052

Mc Laughlin, M., Lu, T., Dimitrijevic, A., & Zeng, F. (2012). Towards a Closed-Loop Cochlear Implant System: Application of Embedded Monitoring of Peripheral and Central Neural

Activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(4), 443–454. https://doi.org/10.1109/TNSRE.2012.2186982

Miller, C. A., Brown, C. J., Abbas, P. J., & Chi, S.-L. (2008). The clinical application of potentials evoked from the peripheral auditory system. *Hearing Research*, *242*(1), 184–197. https://doi.org/10.1016/j.heares.2008.04.005

Müller, N., & Weisz, N. (2012). Lateralized Auditory Cortical Alpha Band Activity and Interregional Connectivity Pattern Reflect Anticipation of Target Sounds. *Cerebral Cortex*, *22*(7), 1604–1613. https://doi.org/10.1093/cercor/bhr232

Nelissen, N., Stokes, M., Nobre, A. C., & Rushworth, M. F. S. (2013). Frontal and Parietal Cortical Interactions with Distributed Visual Representations during Selective Attention and Action Selection. *Journal of Neuroscience*, *33*(42), 16443–16458. https://doi.org/10.1523/JNEUROSCI.2625-13.2013

O'Brien, F., & Cousineau, D. (2014). Representing Error bars in within-subject designs in typical software packages. *The Quantitative Methods for Psychology*, *10*(1), 56–67. https://doi.org/10.20982/tqmp.10.1.p056

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers

into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Prins, N., & Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01250

Ramekers, D., Versnel, H., Strahl, S. B., Smeets, E. M., Klis, S. F. L., & Grolman, W. (2014). Auditory-Nerve Responses to Varied Inter-Phase Gap and Phase Duration of the Electric Pulse Stimulus as Predictors for Neuronal Degeneration. *Journal of the Association for Research in Otolaryngology*, *15*(2), 187–202. https://doi.org/10.1007/s10162-013-0440-x

Salo, E., Salmela, V., Salmi, J., Numminen, J., & Alho, K. (2017). Brain activity associated with selective attention, divided attention and distraction. *Brain Research*, *1664*, 25–36. https://doi.org/10.1016/j.brainres.2017.03.021

Sanchez, G., Lecaignard, F., Otman, A., Maby, E., & Mattout, J. (2016). Active SAmpling Protocol (ASAP) to Optimize Individual Neurocognitive Hypothesis Testing: A BCI-Inspired Dynamic Experimental Design. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00347

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W. R., & Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience*, *22*(11), 2917–2926. https://doi.org/10.1111/j.1460-9568.2005.04482.x

Stypulkowski, P. H., & van den Honert, C. (1984). Physiological properties of the electrically stimulated auditory nerve. I. Compound action potential recordings. *Hearing Research*, *14*(3), 205–223. https://doi.org/10.1016/0378-5955(84)90051-0

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A*, *194*(2), 169–183. https://doi.org/10.1007/s00359-007-0274-2

Tejani, V. D., Abbas, P. J., Brown, C. J., & Woo, J. (2019). An improved method of obtaining electrocochleography recordings from Nucleus Hybrid cochlear implant users. *Hearing Research*, *373*, 113–120. https://doi.org/10.1016/j.heares.2019.01.002

Terreros, G., & Delano, P. H. (2015). Corticofugal modulation of peripheral auditory responses. *Frontiers in Systems Neuroscience*, *9*. https://doi.org/10.3389/fnsys.2015.00134

Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026

Warr, W. B., & Guinan, J. J. (1979). Efferent innervation of the organ of corti: Two separate systems. *Brain Research*, *173*(1), 152–155. https://doi.org/10.1016/0006-8993(79)91104-1

Weise, A., Hartmann, T., Schröger, E., Weisz, N., & Ruhnau, P. (2016). Cross-modal distractors modulate oscillatory alpha power: The neural basis of impaired task performance. *Psychophysiology*, *53*(11), 1651–1659. https://doi.org/10.1111/psyp.12733

Weisz, N., Hartmann, T., Müller, N., & Obleser, J. (2011). Alpha Rhythms in Audition: Cognitive and Clinical Perspectives. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00073

Weisz, N., Müller, N., Jatzev, S., & Bertrand, O. (2014). Oscillatory Alpha Modulations in Right Auditory Regions Reflect the Validity of Acoustic Cues in an Auditory Spatial Attention Task. *Cerebral Cortex*, *24*(10), 2579–2590. https://doi.org/10.1093/cercor/bht113

Winer, J. A. (2006). Decoding the auditory corticofugal systems. *Hearing Research*, *212*(1), 1–8. https://doi.org/10.1016/j.heares.2005.06.014

Wittekindt, A., Kaiser, J., & Abel, C. (2014). Attentional Modulation of the Inner Ear: A Combined

Otoacoustic Emission and EEG Study. *Journal of Neuroscience*, *34*(30), 9995–10002. https://doi.org/10.1523/JNEUROSCI.4861-13.2014

Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *NeuroImage*, *109*, 429–437. https://doi.org/10.1016/j.neuroimage.2014.12.083

Supplemental Information

Methods S1: Details on electrically evoked compound action potentials (ECAPs)

ECAPs were biphasic pulses (anodal polarity of the first pulse phase) with a 40 µs phase duration and an 147 µs interpulse interval. In each participant, the first (i.e. most apical) electrode was used for stimulation and the second for recording. Phase amplitudes and amount of ECAPs measured in each participant were defined between the minimum amplitude given by the electrical hearing threshold and the maximum amplitude given by the maximum loudness of the standard tone (phase amplitude: in steps of 9.45 current units (CU); amount: in steps of one).

Age	Date(s) of implantation	Deafness onset	Reason(s)	CI model(s)	Side (used)	Residual hearing
53	Unknown	Unknown	Unknown	Unknown	Left	No
51	Unknown	Unknown	Unknown	Unknown	Both (left)	No
44	02/2018	Postlingual	Unknown	RONDO 2	Right	Yes
37	04/2018	Postlingual	Stroke	RONDO 2	Right	Yes
52	10/2015, 12/2016	Postlingual	Sudden deafness	SONNET	Both (left)	No
42	01/2018, 01/2019	Postlingual	Perinatal complications	RONDO 2	Both (right)	No
56	01/2020	Postlingual	Hearing loss in higher frequencies	SONNET 2	Right	Yes
38	07/2011, 08/2020	Postlingual	Sudden deafness (left), middle ear inflammation (right)	SONNET 2	Both (left)	Yes
64	04/2019	Postlingual	Unknown	RONDO 2	Left	Yes
57	06/2018, 06/2019	Postlingual	Age-related hearing loss	SONNET 2 (left),	Both (right)	No

				SONNET (right)		
74	12/2002, 08/2012	Postlingual	Ostitis	SONNET 2	Both (left)	No
78	01/2020	Postlingual	Age-related hearing loss	SONNET 2	Left	No
57	08/2014	Postlingual	Sudden deafness	RONDO 2	Right	Yes
70	09/2017	Postlingual	Sudden deafness	SONNET	Right	No
52	06/2014	Postlingual	Age-related hearing loss	SONNET	Left	Yes
55	09/2014, 09/2016	Prelingual	Congenital deafness	SONNET	Both (left)	No

Table S1. *Overview of CI subjects.* Age (in years) is at the time of the experiment session. In the first two participants, only the age and data of the used side was collected.