A direct comparison of voice pitch processing in acoustic and electric hearing

Kurt Steinmetzger a, b, Bastian Meinhardt a, Mark Praetorius b, Martin Andermann a, André Rupp a

a Section of Biomagnetism, Department of Neurology, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany
b Section of Otology and Neurootology, ENT Clinic, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

ARTICLE INFO

Keywords:
Cochlear implants
EEG source reconstructions
fNIRS
Prosody
Single-sided deafness
Speech

ABSTRACT

In single-sided deafness patients fitted with a cochlear implant (CI) in the affected ear and preserved normal hearing in the other ear, acoustic and electric hearing can be directly compared without the need for an external control group. Although poor pitch perception is a crucial limitation when listening through CIs, it remains unclear how exactly the cortical processing of pitch information differs between acoustic and electric hearing. Hence, we separately presented both ears of 20 of these patients with vowel sequences in which the pitch contours were either repetitive or variable, while simultaneously recording functional near-infrared spectroscopy (fNIRS) and EEG data. Overall, the results showed smaller and delayed auditory cortex activity in electric hearing, particularly for the P2 event-related potential component, which appears to reflect the processing of voice pitch information. Both the fNIRS data and EEG source reconstructions furthermore showed that vowel sequences with variable pitch contours evoked additional activity in posterior right auditory cortex in electric but not acoustic hearing. This surprising discrepancy demonstrates, firstly, that the acoustic detail transmitted by CIs is sufficient to distinguish between speech sounds that only vary regarding their pitch information. Secondly, the absence of a condition difference when stimulating the normal-hearing ears suggests a saturation of cortical activity levels following unilateral deafness. Taken together, these results provide strong evidence in favour of using CIs in this patient group.

1. Introduction

A major challenge when studying functional brain activity in clinical populations is that the control groups required to evaluate the results can only be matched for a limited set of parameters. Here, we present data from a group of adults with single-sided deafness (SSD), all of which were fitted with cochlear implants (CIs) in the deafened ear but had preserved normal acoustic hearing in the other ear, resulting in the rare scenario that each participant acted as their own control subject. Separately stimulating both ears allowed for a direct comparison of acoustic hearing and CI-based electric hearing.

CIs have proven very successful in restoring hearing after profound sensorineural hearing loss (Macherey and Carlyon, 2014; Wilson and Dorman, 2008). Yet, the electrical components of the devices have hindered the study of brain activity in CI users, particularly the use of functional MRI and MEG. EEG data obtained from CI users, on the other hand, are severely affected by electrical artefacts, and a substantial part of the literature has focussed on ways to remove them (Friesen and Picton, 2010; Gilley et al., 2006; Viola et al., 2012; Viola et al., 2011). Therefore, PET has been applied extensively in CI research (Anderson et al., 2017a; Strelnikov et al., 2015), despite its inconvenience. More recently, however, functional near-infrared spectroscopy (fNIRS) has been identified as an ideal CI-compatible imaging method (Anderson et al., 2017a; Pinti et al., 2018).

The majority of PET studies with CI users focussed on how the processing of speech is affected by deafness-induced plasticity effects in auditory and visual cortex and how these relate to CI-based speech intelligibility (Coez et al., 2008; Giraud et al., 2001a; Giraud et al., 2017a).
neural activity in CI-based pitch processing cannot be inferred from the distribution of cortical activity in detail nor did it include source-level ERP analyses.

Previously, we tested young NH listeners with the same paradigm and methodology to obtain results the present data can be compared with and to assess the congruence of the two imaging methods (Steinmetzger et al., 2022). Both the fNIRS topographies and the ERP source reconstructions showed strongly right-lateralised activity in superior temporal areas for both types of vowel sequences. Sequences with variable prosodic contours showed strongly right-lateralised activity in superior temporal areas for CI users and normal-hearing (NH) controls are lacking. This is particularly striking with regard to pitch perception, which is very limited when listening through a CI and a crucial factor underlying the difficulties CI users encounter when listening to speech and music, especially in background noise (Green et al., 2004; Oxenham, 2008; Steinmetzger and Rosen, 2016). The only pitch cues available to CI users are the periodic fluctuations of the temporal envelope of the acoustic input and hence their ability to identify the voice pitch contours that constitute a major part of prosody is strongly reduced (Chatterjee and Peng, 2008; Everhardt et al., 2020).

Here, the participants were presented with continuous sequences of natural vowels in which the pitch contours were either the same throughout or varied between the individual vowels, while controlling for spectro-temporal differences between the two conditions. This subtle but linguistically relevant acoustic contrast allowed us to compare the processing of voice pitch changes in acoustic and electric hearing. Concurrent fNIRS and EEG data were obtained while stimulating each ear separately to study the evoked haemodynamic and electrophysiological activity. Furthermore, source reconstructions of the resulting event-related potentials (ERPs) were computed to investigate the interrelation of both types of cortical activity in detail. To date, there is only one concurrent fNIRS-EEG study with CI users (Chen et al., 2017b), which reported lower adaptation rates to visual and auditory stimuli compared to NH controls, but this study neither examined the distribution of cortical activity in detail nor did it include source-level ERP analyses.

More generally, both diminished haemodynamic cortical activity in auditory areas (Coez et al., 2008; Giraud et al., 2000) as well as smaller and delayed auditory ERPs (Agrawal et al., 2013; Kelly et al., 2005; Sandmann et al., 2009; Sandmann et al., 2015; Viola et al., 2011; Wagner et al., 2017) have been observed in CI users relative to NH controls using different types of auditory stimuli. These effects were shown to increase with the duration of deafness before implantation (Green et al., 2005; Kelly et al., 2005; Sandmann et al., 2010; Viola et al., 2011) and possible explanations for them include the lower number of activated neurons as well as a reduction of synchronised neural activity caused by the limited sensory input provided by CIs (Agrawal et al., 2013; Sandmann et al., 2009). However, these differences between acoustic and electric hearing have never been shown in the same group of subjects. Similar effects in the current sample of unilateral CI users with preserved contralateral acoustic hearing, particularly regarding the delayed ERPs, would thus imply temporally disparate neural responses despite several years of experience with bimodal hearing.

2. Material and methods

2.1. Participants

Twenty participants with SSD, i.e., unilateral CI users with preserved normal acoustic hearing in the contralateral ear, took part in this experiment. All subjects were post-lingually deafened German native speakers. According to the current WHO (2021), three subjects had a moderate hearing loss in the contralateral ear (35–50 dB average hearing level from 0.5 to 4 kHz; subjects 5, 17 & 20). For all other subjects, the audiometric thresholds only showed some age-typical high-frequency hearing loss at the time of implantation (Fig. 1). None of the participants wore a hearing aid in the contralateral ear or were receiving clinical treatment for hearing loss in the non-implanted ear when the experiment was conducted, indicating that their audiometric thresholds have not deteriorated significantly. Detailed information regarding the participants is provided in Table 1. The words correct scores for the CI ears were determined with the Freiburg monosyllabic speech intelligibility test at a presentation level of 65 dB SPL (Hahlbrock, 1953). Written consent was obtained prior to the experiment. The study was approved by the local research ethics committee (Medical Faculty, University of Heidelberg) and was conducted in accordance with the Declaration of Helsinki.

2.2. Stimuli

The materials used in this study were identical to those employed in a previous fNIRS-EEG experiment with young NH listeners (Steinmetzger et al., 2022), where the stimulus construction is described in more detail.
Waveforms and narrow-band spectrograms of an example vowel synthesised with the five different prosodic contours are shown in Fig. 2 A.

The individual vowels were presented in continuous blocks (20 vowels/16 s) followed by pauses (16 s), as shown in Fig. 2B. The experimental design consisted of two conditions: In the FIXED PROSODY condition, all vowels within a block had the same prosodic contour (flat, rising straight, falling straight, rising curved, or falling curved). In the VARIABLE PROSODY condition, the contours varied between the vowels within each block (rising, falling, straight, curved, or a mixture of all five contour types). The five different contour types were intended to represent a set of typical, easily distinguishable prosodic contours found across languages. The range of prosodic contours, vowels, and F0s was primarily included to

The experimental design was identical to that described in Steinmetzger et al. (2022), except that each subject completed the experiment twice, once for the NH ear and once for the CI ear. The order of the two sessions was counterbalanced and each subject went through the whole testing procedure on the same day.

The individual vowels were presented in continuous blocks (20 vowels/16 s) followed by pauses (16–20 s), as shown in Fig. 2B. The experiment consisted of two conditions: In the FIXED PROSODY condition, all vowels within a block had the same prosodic contour (flat, rising straight, falling straight, rising curved, or falling curved). In the VARIABLE PROSODY condition, the contours varied between the vowels within each block (rising, falling, straight, curved, or a mixture of all five contour types). The five different contour types were intended to represent a set of typical, easily distinguishable prosodic contours found across languages. The range of prosodic contours, vowels, and F0s was primarily included to

ensure the generalisability of the results and to make the experiment less monotonous. Every participant was presented with 50 blocks per condition, in random order. Each session of the experiment thus consisted of 100 stimulus blocks, amounting to a duration of about 57 mins. As the EEG data were analysed relative to the onset of the individual vowels in each block, this design resulted in 2000 trials per session.

Table 1
Subject information

| Subject | Age | Sex | CI ear | Duration SSD (years,months) | Duration CI use (years,months) | Aetiology of deafness | CI & processor type / strategy | Words correct CI ear (%) |
|---------|-----|-----|--------|-----------------------------|-------------------------------|----------------------|-------------------------------|--------------------------|
| 1       | 58  | m   | l      | 23                          | 5.5                           | Intracochlear schwannoma | FLEX28 & OPUS2 / FS4-p       | 60                       |
| 2       | 61  | f   | r      | 6                           | 5.6                           | Acoustic neuroma         | FLEX28 & OPUS2 / FS4          | 65                       |
| 3       | 59  | f   | l      | 1                           | 2.2                           | Sudden hearing loss      | HiRes90K & Naida Q90 / HiRes Optima-S | 45                       |
| 4       | 66  | f   | r      | 26                          | 2.6                           | Sudden hearing loss      | FLEX28 & RONDO / FS4-p       | 65                       |
| 5       | 66  | f   | l      | 22                          | 5.6                           | Sudden hearing loss      | FLEX28 & OPUS2 / FS4          | 10                       |
| 6       | 67  | f   | l      | 1                           | 5.2                           | Sudden hearing loss      | CONCERTO medium & OPUS2       | 45                       |
| 7       | 66  | m   | l      | 1                           | 6.1                           | Sudden hearing loss      | C422 & CP10 / ACE            | 70                       |
| 8       | 55  | f   | r      | 39                          | 6.1                           | Mumps                   | FLEX28 & OPUS2 / FS4          | 55                       |
| 9       | 50  | f   | l      | 1                           | 5.9                           | Sudden hearing loss      | FLEX28 & OPUS2 / FS4-p       | 45                       |
| 10      | 44  | f   | r      | 2                           | 4.4                           | Otosclerosis             | CI522 & CP90 / ACE           | 55                       |
| 11      | 67  | f   | l      | 1                           | 6.7                           | Sudden hearing loss      | C422 & CP10 / ACE            | 35                       |
| 12      | 42  | f   | r      | 1                           | 5.3                           | Sudden hearing loss      | HiRes90K & Naida Q90 / HiRes Optima-S | 80                       |
| 13      | 63  | f   | l      | 3                           | 3.7                           | Sudden hearing loss      | FLEX28 & RONDO / FS4-p       | 55                       |
| 14      | 77  | f   | r      | 13                          | 2.10                          | Sudden hearing loss      | FLEX28 & SONNET / FS4        | 30                       |
| 15      | 60  | m   | r      | 1                           | 3.7                           | Sudden hearing loss      | FLEX28 & RONDO / FS4-p       | 35                       |
| 16      | 78  | f   | r      | 1                           | 5.0                           | Sudden hearing loss      | FLEX28 & SONNET / FS4        | 35                       |
| 17      | 70  | m   | r      | 1                           | 2.1                           | Sudden hearing loss      | HiRes Ultra & Naida Q90 / HiRes Optima-S | 70                       |
| 18      | 26  | f   | r      | 1                           | 3.4                           | Meningitis / Temporal bone fracture | FLEX28 & SONNET / FS4 | 80                       |
| 19      | 66  | m   | r      | 30                          | 1.4                           | Sudden hearing loss      | FLEX28 & ROND02 / FS4-p      | 55                       |
| 20      | 58  | m   | l      | 20                          | 4.1                           | Unknown                  | HiRes90K & Naida Q70 / HiRes Optima-S | 90                       |

Ø = 60 \( \frac{f}{12} = \frac{r}{14} \), \( \theta = 10 (12) \), \( \theta = 4.3 (1.7) \), Sudden hearing loss = 14

MED-EL = 13, \( \theta = 54 (19.6) \)

a Group means (standard deviations) and overall counts, respectively, are provided in the bottom row.
Both ears were tested using free-field acoustic stimulation, for which the stimuli were converted with 24-bit resolution at a sampling rate of 48 kHz using an ADI-8 DS sound card (RME, Haimhausen, Germany) and presented via an Adam A7x speaker (Adam Audio, Berlin, Germany). The speaker was placed directly in front of the listener, ~1.5 m away and at ear level. The presentation level was set to 70 dB SPL using a sound level meter (Bruel & Kjær, type 2235; Nærum, Denmark) placed at the position of the subject's head. When testing the CI ear, the normal ear was blocked with a combination of earplug (3M Series 1100; Maplewood, MN, USA; ~37 dB attenuation) and over-ear hearing protection (3M Peltor Optime III H540A; ~35 dB attenuation). To test the NH ear, the external CI processor was taken off (Fig. 2C). The experiment took place in a sound attenuating and electrically shielded room, with the participant sitting in a comfortable reclining chair during data acquisition. To minimise ambient light from interfering with the recordings, the participants wore an overcap and the room light was dimmed to the lowest level. There was no behavioural task, but pauses were inserted about every 15 mins to ensure the vigilance of the subjects.

2.4. fNIRS recording and analysis

The fNIRS methodology used in the current study is largely the same as in Steinmetzger et al. (2022), where it is described in greater detail. fNIRS signals were recorded with a continuous-wave NIRSScout 16x16 system (NIRx Medizintechnik, Berlin, Germany) at a sampling rate of 7.8125 Hz. Eight source optodes and eight detector optodes were placed symmetrically over each hemisphere by mounting them on an EEG cap with holes near the electrode sites P7 and P8 to accommodate the CI coils (EasyCap, Herrsching, Germany). The source optodes emitted infrared light with wavelengths of 760 and 850 nm. The chosen optode layout was devised to optimally cover the auditory cortex and associated areas. As the CI speech processors of most subjects hindered a stable scalp contact of optodes in the inferior temporal region, the corresponding channels (left: ch# 9 & 10; right: ch# 31 & 32) were excluded from the analyses. This resulted in 20 measurement channels per hemisphere, of which 18 had a standard source-to-detector distance of about 30 mm, while the remaining 2 had a shorter spacing of about 15 mm. The optode and reference positions for each individual subject were digitised with a Polhemus 3SPACE ISOTRAK II system (Colchester, VT, USA) before the recordings.

The raw data were pre-processed using the HOMER2 toolbox (version 2.8; Huppert et al. 2009) and custom MATLAB code. The raw light intensity signals were first converted to optical density values and then corrected for motion artefacts. A kurtosis-based wavelet algorithm with a threshold value of 3.3 was used to identify and correct motion artefacts (Ciarelli et al., 2015). Measurement channels with poor signal quality were then identified by computing their scalp coupling index (SCI; Pollonini et al. 2014) and excluded from further analysis if the SCI value was smaller than 0.6 (NH data sets: mean = 0.35 channels/subject, max 4 per subject; CI data sets: mean = 0.4, max = 3). Next, the motion-corrected signals of the remaining channels were band-pass filtered between 0.01 and 0.5 Hz to isolate the task-related neural activity, and subsequently converted to concentration values based on the modified Beer-Lambert law (Scholkmann et al., 2014). The differential path length factors required for the conversion were determined based on the wavelengths of the light and the age of the subject (Scholkmann and Wolf, 2013).

The pre-processed data were statistically evaluated and topographically visualised with SPM-NIRS (version r3; Tak et al. 2016). The optode positions of each subject were first transformed from subject space to MNI space, after which they were probabilistically rendered onto the ICBM-152 cortical template surface (Singh et al., 2005). The signals were then temporally smoothed using the shape of the canonical haemodynamic response function (HRF, ‘pre-colouring’) to avoid autocorrelation issues (Worsley and Friston, 1995). The data of the individual subjects were statistically modelled by convolving the continuous signals obtained from each long channel with 16-s statistical parametric mapping (SPM) double-gamma functions representing the stimulus blocks. The oxygenated (HbO) and de-oxygenated (HbR) haemoglobin data were modelled with positive and negative HRFs, respectively. To allow the time course of the measured concentration changes to vary slightly, the temporal and spatial derivatives of the canonical HRF were included as additional regressors (Plichta et al., 2007). Furthermore, the signals from the four short channels were subjected to a principal component analysis, using the first component as an additional nuisance regressor to remove the so-called global scalp-haemodynamic component (Sato et al., 2016), i.e., the superficial signal component which is thought to not reflect cortical activity.

After estimating the HbO and HbR general linear models (GLMs) for each subject, contrast vectors were defined to assess the functional activations. When comparing the activity across ears or conditions, the regressors of interest were set to 1 and −1, respectively, while the regressors representing the derivatives and the global scalp component were set to 0 to statistically control for their effects. Likewise, when evaluating the activity within conditions, the respective regressors were set to 1, whereas all other regressors were set to 0. Group-level statistics of each contrast were then computed for each long channel by assessing the subject-level beta weights using one-sided t-tests.

A customised version of the SPM-fNIRS plotting routine was devised to visualise the optode and channel positions as well as the functional activations. The optode locations as well as the resulting channel positions and distances for the individual subjects are shown in Supplementary Fig. 1. Two auditory regions of interest (ROIs) were defined a priori, each comprising four channels positioned along the STC in both hemispheres (left ROI: ch# 5, 12, 14, and 16; right ROI: ch# 27, 34, 36, and 38). In the plots showing the functional activations, the short channels were omitted as they are assumed to not reflect any cortical responses. In the HRF plots, the subject-level waveforms were baseline corrected by subtracting the mean amplitude from −2−0 s around block onset from each sample point. The HRFs are shown both after the preprocessing (‘Total HRF’) as well as after regressing out the contribution of the short channels and pre-colouring the signals (‘Cortical HRF’), to illustrate the effect of removing the non-cortical signal component. Additionally, the β-weighted canonical HRFs (‘Modelled HRF’) are shown to demonstrate how well the cortical HRFs can be explained by the GLM.

2.5. EEG recording and analysis

Continuous EEG signals were recorded using a 64-channel BrainVista actiChamp system (Brain Products, Gilching, Germany). CI-compatible custom EEG caps with holes for the transmitter coils at electrode positions P7 and P8 were used. Apart from this deviation, the 60 scalp electrodes were arranged according to the extended international 10–20 system. Four additional electrodes were placed around the eyes to record vertical and horizontal eye movements. The EEG data were recorded with an initial sampling rate of 500 Hz, an online anti-aliasing low-pass filter with a cut-off frequency of 140 Hz, and were referenced to the right mastoid. The electrode positions of each subject were digitized with a Polhemus 3SPACE ISOTRAK II system before the experiment too.

The raw data were pre-processed offline using FieldTrip (version 20180924; Oostenveld et al. 2011) and custom MATLAB code. The continuous waveforms were first segmented into epochs ranging from ~0.3–0.9 s relative to vowel onset. Next, linear trends and the DC component were removed by subtracting a 1st-order polynomial and the epochs were low-pass filtered at 15 Hz. The epochs were then referenced to the mean of both mastoids and down-sampled to 250 Hz. After visually identifying and excluding bad channels (NH data sets: mean = 1.25 channels/subject, max 7 per subject; CI data sets: mean = 1.8, max = 12), the data were decomposed into 20 principal components to detect and eliminate eye artefacts.
Pronounced CI artefacts were only evident in a subset of participants (8/20), for whom the principal component analysis was performed a second time to further reduce them (Schierholz et al., 2017; Schierholz et al., 2015). As CI artefacts are particularly large following sound onset and offset (Gilley et al., 2006; Viola et al., 2012), the use of continuous stimulus sequences appears to be the main reason for the relatively small artefact sizes in the current data (Pantev et al., 2006).

After the four eye electrodes were removed from the data, epochs in which the amplitudes between −0.2–0.8 s around stimulus onset exceeded ± 60 µV or the z-transformed amplitudes differed by more than 15 standard deviations from the mean of all channels were excluded from further processing. On average, 88% of the trials (1756/2000) passed the rejection procedure for the NH data sets and 87% for the CI data sets (1745/2000). In one subject, an unstable contact of the ground electrode led to data loss, so that only 42% (841/2000) of the NH trials and 36% (718/2000) of the CI trials remained after the rejection procedure. Lastly, bad channels were interpolated using the weighted average of the neighbouring channels, the data were re-referenced to the average of all 60 channels, and baseline corrected by subtracting the mean amplitude from −0.1–0 s before stimulus onset.

The pre-processed ERP data were then exported to BESA (version 5.2; BESA GmbH, Gräfeling, Germany; Scherg 1990) for source analysis. Based on the grand-average ERPs of the NH ears pooled across all trials, a source model with one current dipole per hemisphere was devised. A four-shell ellipsoidal head model was used. The time window used for fitting the model was centred around the P2 peak (216–372 ms). The dipole locations were constrained to be symmetrical across hemispheres, whereas the dipole orientations were estimated without constraints. For the grand-average ERPs obtained via the NH ears, this source model explained 94.96% of the variance in the data. To account for the lower signal-to-noise ratio of the data and to facilitate the comparability of the results between the ears, the same model was then applied without any alterations to the ERPs obtained via the CI ears too.

3. Results

3.1. fNIRS results

The group-level fNIRS HbR topographies are shown in Fig. 3A. The upper and middle rows depict the functional activations of each channel for the NH and CI ears compared to baseline level, whereas the lower row shows where the activity was stronger when the participants listened through the NH ears compared to the CI ears. The auditory cortex activity in these three contrasts was statistically evaluated by separately testing each channel included in the two auditory ROIs (left: ch# 5, 12, 14, & 16; right: ch# 27, 34, 36, & 38). The false discovery rate (FDR) across the ROI channels was controlled using the Benjamini-Hochberg procedure.

When the NH ears were stimulated, all eight ROI channels showed significant activity (averaged Cohen’s d for right ROI = 1.08, left ROI = 0.74), whereas only two channels in the left ROI (ch# 5 & 16; average d = 0.55) reached significance level when the CI ears received auditory input (one-sample t-tests against 0, thresholded at $p_{(FDR)} < 0.05$). The comparison of both ears revealed significantly greater activity for three channels along the right STC (ch# 27, 34, & 36; right ROI: $d = 0.74$) and one channel near the left primary auditory cortex (ch# 12; left ROI: $d = 0.44$) for the NH ears (paired-samples t-tests, thresholded at $p_{(FDR)} < 0.05$). Possible hemispheric lateralisation effects were then assessed by averaging the beta weights of the four channels in each ROI and comparing them across hemispheres. Paired t-tests indicated no hemispheric asymmetries for the NH ($t_{(19)} = 0.31, p = 0.759$) as well as the CI ears ($t_{(19)} = -0.54, p = 0.594$). To complement the topographical results, the corresponding HRFs are shown in Fig. 3B, separately for both ROIs. As for the topographies, the HRF amplitudes were markedly larger and more sustained for the NH ears.

We then compared the HbR topographies across the two stimulus conditions (Fig. 4, full results provided in Supplementary Fig. 2). For the NH ears, none of the channels in the two auditory ROIs showed greater activity in the variable prosody condition (paired-samples t-tests, thresholded at $p_{(FDR)} < 0.05$). Even when not correcting for multiple
comparisons, the trend in the posterior part of the right STC (ch# 38) did not reach significance level (t19) = 1.64, p = 0.059). For the CI ears, in contrast, two channels in the right ROI (ch# 34 & 36) as well as an adjacent channel (ch# 37) showed significantly greater activity in the VARIABLE PROSODY condition (t ≥ 2.39, p < 0.014).

The corresponding fNIRS HbO results are provided in Supplementary Fig. 3. As in a previous auditory perception study with young NH listeners (Steinmetzger et al., 2020), the HbO topographies showed a prominent bilateral gradient with increasingly more negative effects in anterior-inferior direction. Rather than indicating a deactivation of fronto-temporal cortical areas, the fNIRS and EEG analyses in that paper demonstrated that this effect was associated with larger auditory ERPs and increased attention levels in the subjects who showed this pattern. Moreover, the negative HRFs observed here and in the earlier paper were found to lag behind the canonical HRF model considerably, suggesting that this effect does not reflect cortical activity but may be due to increased sympathetic nervous system activity (Kirilina et al., 2012; Tachtsidis and Scholkmann, 2016). However, as this gradient masks the underlying cortical activity in auditory cortex, we are focussing on the HbR data in the current paper.

3.2. EEG results

The sensor-level ERPs across all stimuli are shown in Fig. 5A, separately for both ears. The data show the pooled activity of a set of 16 electrodes in the fronto-central scalp region, where auditory ERPs typically have the largest amplitudes. The waveforms show a comparable morphology for both ears, with the characteristic transient ERP components (P1, N1, and P2) followed by a sustained potential (SP). A pronounced amplitude difference between the two ears is evident for the P2, which was markedly larger when the NH ears were stimulated. Permutation-based (n = 10,000) paired-samples t-tests for each time point during the stimulus window (0–800 ms) confirmed this observation, as differences between the two ears were confined to this period (212–276 ms, p < 0.001). Additionally, the N1 (t19) = -4.59, p < 0.001, d = 1.03; 176/198 ms) and the P2 (t19) = -4.36, p < 0.001, d = 0.97; 300/330 ms) peaked significantly earlier when the NH ears received auditory stimulation. The latency windows from which the respective peaks were extracted are based on the root-mean-square (RMS) waveforms of the grand-average ERPs across all channels. The scalp maps for each component time window and ear are shown on the right of Fig. 5A and exhibit similar distributions for both ears.

The corresponding source-level ERPs are shown in Fig. 5B. The cortical generators of the P2 were modelled using two dipoles located in the left and right planum temporale (MNI coordinates: ±43, −35, 13), posterior to primary auditory cortex. When averaging across both dipoles, the resulting source waveforms were very similar to the sensor-level ERPs described above: The largest amplitude differences were again observed for the P2 (208–296 ms, p < 0.001) and shorter latencies were evident for the N1 (t19) = -5.79, p < 0.001, d = 1.29; 171/196 ms), the P2 (t19) = -3.30, p = 0.002, d = 0.74; 292/330 ms), and also the SP (t19) = -2.38, p = 0.014, d = 0.53; 500/542 ms) when the NH ears were stimulated. As shown in Fig. 5C, the dipole source modelling also allowed a separate analysis of the ERPs originating from both hemispheres. To aid the spatial comparison with the fNIRS results, the upper part of the figure shows the dipole locations after projecting them onto the surface of the template cortex used in the fNIRS analyses. The projected dipole positions corresponded to locations at the posterior end of the auditory fNIRS ROIs. The distance from the original positions to the scalp surface was about 41 mm (MNI coordinates: ±84, −35, 13).

The source waveforms revealed, firstly, that the P2 amplitude was significantly larger (p < 0.001) for the NH ears in both hemispheres. Additionally, the P2 emerging from the right hemisphere was found to have a longer latency for the NH (t19) = 3.32, p = 0.004, d = 0.74; 273/314 ms) as well as the CI ears (t19) = 2.30, p = 0.033, d = 0.51; 320/349 ms), but the comparison of the hemispheric differences across ears revealed no hemisphere*ear interaction (t19) = 0.70, p = 0.495). The dipole orientations also reflect this hemispheric difference, as the right dipole shows a distinctly more lateral orientation (Fig. 5B). Note that the apparent N1 amplitude difference across hemispheres is disregarded here as the dipole model was devised to localise the P2 rather than the N1, and the right dipole may thus only have captured parts of the N1.

Next, we examined whether the source-level ERP amplitudes were larger in the VARIABLE PROSODY condition than in the fixed PROSODY condition (Fig. 5D). Surprisingly, no such effect was evident for the NH ears, neither for the P2 nor any other component. For the CI ears, in contrast, a significantly larger P2 amplitude (304–344 ms, p < 0.05) was observed in the right hemisphere. A similar pattern of results was also observed at the scalp level (Supplementary Fig. 4). Moreover, the source-level ERPs revealed an ear*condition interaction for a similar time window (316–336 ms, p < 0.05), reflecting that the condition difference for the P2 amplitudes was significantly larger for the CI ears.

3.3. Relation of fNIRS and EEG data with CI-based speech intelligibility

In a next step it was evaluated if the speech intelligibility scores that the participants achieved when listening through their CI ears (cf. Table 1) were reflected in the fNIRS and EEG data. When comparing the subjects with scores above and below the group median (55% words correct, n = 8 per subgroup), the fNIRS HbR data averaged across stimulus conditions indicated significantly greater activity in the posterior part of right auditory cortex (ch# 36; t18 = 2.61, p = 0.01, d = 1.30; independent-samples t-test) for the high intelligibility subgroup (Fig. 6A). Likewise, a significant negative correlation of the mean HbR amplitudes of all 20 subjects during the stimulus period (0–16 s) and their speech intelligibility scores was observed for channel 36 (r18 = -0.54, p = 0.007; Fig. 6A), while no other channel in the two auditory ROIs showed a similar effect (p ≥ 0.138).

To further evaluate whether the gradient in the fNIRS HbO data reflects better auditory processing, the mean HbO amplitudes during the stimulus blocks (0–16 s) were then averaged across all channels in the two auditory ROIs and correlated with the CI speech intelligibility...
scores. As shown in Fig. 6B, the HbO amplitudes indeed exhibited a significant negative correlation with the percentage of correctly perceived words in the Freiburg monosyllabic speech test ($r(18) = -0.48$, $p = 0.016$).

To complement the intelligibility effect observed in the fNIRS HbR data, it was furthermore tested whether the source-level ERPs of the dipole in the right hemisphere also reflect the CI-based speech intelligibility scores (Fig. 6C). For the right dipole, located near fNIRS channel 36, the transient ERPs in the high intelligibility subgroup were indeed more pronounced and significantly larger amplitudes were observed for the N1 ($p < 0.05$). Similarly, the subjects’ N1 amplitudes (175–240 ms) in the right hemisphere and their intelligibility scores were negatively correlated ($r(18) = -0.46$, $p = 0.02$).

4. Discussion

4.1. Smaller and delayed cortical responses in CI-based electric hearing

In the present study, auditory cortex activity evoked by acoustic and electric stimulation was directly compared in a group of unilateral adult CI users with SSD. A key finding is that cortical responses were larger when the NH ears received auditory stimulation, as evidenced by the fNIRS as well as the EEG data. The EEG results furthermore showed that the ERPs had significantly shorter latencies when the participants listened via their NH ears.

The fNIRS data revealed smaller responses for the CI ears, representing the first BOLD-based evidence of a difference between acoustic and electric hearing in the same group of participants. Specifically, the fNIRS HbR results indicated larger cortical activity along the right STC and near left primary auditory cortex for the NH ears. In contrast to a previous fNIRS-EEG study (Steinmetzger et al., 2022), in which young NH listeners tested with the same paradigm showed strongly right-lateralised activity, stimulation of both the NH and CI ears here led to relatively symmetric distributions of activity in auditory areas. For the CI ears, there was even a slight trend towards greater activity in the left hemisphere. With current CI systems, the access to spectral information is severely limited, whereas temporal envelope modulations are conveyed fairly accurately (Chatterjee and Peng, 2008; Macherey and Carlyon, 2014; Steinmetzger and Rosen, 2018). Models of auditory processing propose that the temporal resolution of the left auditory cortex is higher compared to the right one, which is in turn characterised by a greater spectral resolution (Poeppel, 2003; Zatorre et al., 2002). Accordingly, the trend for greater activity in the left hemisphere may reflect that the auditory input received via the CI ears contains more temporal than spectral information.

The corresponding ERP results showed that greater P2 amplitudes in response to stimulation of the NH ears seemed to underly the larger BOLD responses, suggesting that the P2 may be critically involved in the processing of voice pitch contours. Although the dipole source model of the P2 demonstrated that this effect was evident in both hemispheres,
8

Fig. 6. CI-based speech intelligibility. A) fNIRS HbR topography of the contrast between subjects with high and low CI-based speech intelligibility scores. The latter subgroup showed significantly larger activity in the posterior part of the right auditory cortex (ch# 36). Furthermore, the HbR amplitudes at channel 36 and the intelligibility scores were negatively correlated. The words correct scores in the scatter plot are taken from Table 1. B) The fNIRS HbO amplitudes and the intelligibility scores were negatively correlated. The words correct latter subgroup showed significantly larger activity in the posterior part of the right hemisphere. This hemispheric difference is also reflected in the HbR amplitudes at channel 36 across all auditory ROI channels were also negatively correlated with speech intelligibility. C) The source-level ERPs of the dipole in the right hemisphere for right auditory cortex (ch# 36). Furthermore, the HbR amplitudes at channel 36 across all auditory ROI channels were also negatively correlated with speech intelligibility scores in the scatter plot are taken from Table 1.
The tracking of pitch changes in normal hearing has been associated with right-lateralised activity in superior temporal regions anterior to primary auditory cortex (Johnsrude et al., 2000; Poeppel, 2003; Zatorre and Belin, 2001), even if the stimulus materials only contained temporal pitch cues (Patterson et al., 2002). Consequently, the lack of additional activity in those areas in the variable prosody condition suggests that even though the participants were able to differentiate between the two stimulus conditions, they relied on timbral cues rather than the pitch contours to do so (Oxenham, 2008). Nonetheless, this finding implies that the sensory input received via the implant contains enough information to supplement the contralateral normal ear, arguing in favour of fitting a CI to the deafened ear in unilateral deaf subjects.

Furthermore, the amount of activity in posterior right auditory areas coincided with speech intelligibility, as indicated by greater NIIRS HBR responses for channel 36 and increased N1 amplitudes for the right-hemisphere dipole. Thus, the same region that showed differential responses between the two stimulus conditions also reflected the participants’ speech intelligibility scores. As spectral information is preferentially processed in right auditory areas, this finding may indicate a greater spectral resolution of the electric auditory input in subjects with high speech intelligibility.

5. Conclusions

The present study compared the cortical processing of voice pitch information in electric and acoustic hearing in a group of experienced unilateral CI users with contralaterally preserved acoustic hearing. Although poor transmission of pitch cues is a crucial factor underlying the difficulties CI user experience in noisy acoustic environments, it has remained unclear how the cortical processing of pitch information differs across both types of hearing.

The current NIIRS results showed that the pitch contours in natural vowels elicited bilateral activity in auditory cortex in electric as well as acoustic hearing, but that this activity was significantly smaller when the CI ears were stimulated. Coherent with the NIIRS data, the simultaneously obtained EEG data exhibited smaller and delayed ERPs in CI-based hearing. This pattern was particularly pronounced for the P2, which thus appears to reflect the processing of voice pitch information. Additionally, both the NIIRS and EEG results showed that vowel sequences with variable rather than repeating pitch contours evoked additional activity in posterior right auditory cortex in electric but not acoustic hearing. This surprising finding demonstrates that the acoustic detail transmitted by CIs is sufficient to distinguish between speech sounds that only vary regarding their pitch information, while the functionally unspecific activity observed for the NH ears suggests a saturation of cortical activity levels.

Furthermore, when stimulating the CI ears, activity in this region was larger in subjects with higher CI-based speech intelligibility. This shows that the voice-pitched cortical activity patterns obtained here are of diagnostic relevance, as they allow for a better understanding of the performance differences between CI users.

Funding

This work was supported by the Dietmar Hopp Stiftung (Grant No 2301 1239).

CRediT authorship contribution statement

Kurt Steinmetzger: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. Bastian Meinhardt: Investigation, Data curation, Project administration. Mark Praetorius: Conceptualization, Funding acquisition. Martin Andermann: Conceptualization, Funding acquisition, Writing – review & editing. André Rupp: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The stimuli, NIIRS and EEG data, as well as the code used to process the data are all available at https://doi.org/10.17605/osf.io/9v6em.

Acknowledgements

We would like to thank Elisabeth Munk and Madhuri Sharma Rao for their help with recruiting the participants.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.103188.

References

Agrawal, D., Thorne, J., Vila, F., Timm, L., Debenar, S., Büchner, A., Dengler, R., Witte, J., Rothg, M., 2013. Electrophysiological responses to emotional prosody perception in cochlear implant users. NeuroImage: Clinical 2, 229–238. https://doi.org/10.1016/j.nicl.2013.01.001.

Anderson, C.A., Lazard, D.S., Hartley, D.E., 2017a. Plasticity in bilateral superior temporal cortex: effects of deafness and cochlear implantation on auditory and visual speech processing. Hearing Res. 343, 138–149. https://doi.org/10.1016/j.heares.2016.07.013.

Anderson, C.A., Wiggins, I.M., Kitterick, P.T., Hartley, D.E., 2017b. Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. Proc. Natl. Acad. Sci. USA 114, 10256–10261. https://doi.org/10.1073/pnas.1704751114.

Bisconti, S., Shulkin, M., Hu, X., Banaru, G.J., Kileny, P.R., Kovelman, I., 2016. Functional near-infrared spectroscopy brain imaging investigation of phonological awareness and passage comprehension abilities in adult recipients of cochlear implants. J. Speech. Lang. Hear. Res. 59, 239–253. https://doi.org/10.1044/2015_JSLHR-L-14-0278.

Boemio, A., Fromm, S., Braun, A., Poeppl, D., 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat. Neurosci. 8, 389–395. https://doi.org/10.1038/nn1409.

Bonitz, H., Kopp, B., Büchner, A., Lanner, T., Lyekyll, B., Finke, M., 2018. Event-related neuronal responses to acoustic novelty in single-sided deaf cochlear implant users: Initial findings. Clin. Neurophysiol. 129, 133–142. https://doi.org/10.1016/j.clinph.2017.10.025.

Chatterjee, M., Peng, S.-C., 2008. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. Hearing Res. 235, 143–156. https://doi.org/10.1016/j.heares.2007.11.004.

Chen, L.-C., Sandmann, P., Thorne, J.D., Bleichner, M.G., Debenar, S., 2016. Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with NIIRS. Neural Plast. 2016, 4382656. https://doi.org/10.1155/2016/4382656.

Chen, L.-C., Puschmann, S., Debenar, S., 2017a. Increased cross-modal functional connectivity in cochlear implant users. Sci. Rep. 7, 10043. https://doi.org/10.1038/s41598-017-10792-2.

Chen, L.-C., Strohpl, M., Schönwiesner, M., Debenar, S., 2017b. Enhanced visual adaptation in cochlear implant users revealed by concurrent EEG-NIIRS. NeuroImage 146, 600–608. https://doi.org/10.1016/j.neuroimage.2016.09.023.

Chiarelli, A.M., Maclln, E.L., Fabiani, M., Gratton, G., 2015. A kurtosis-based wavelet algorithm for motion artifact correction of NIIRS data. NeuroImage 112, 128–137. https://doi.org/10.1016/j.neuroimage.2015.02.057.

Coen, A., Zihovsic, M., Ferraty, E., Bouscaur, D., Monier, L., Ambert-Dahan, E., Bizugat, E., Sylota, A., Samson, Y., Sterkers, O., 2008. Cochlear implant benefits in deafness rehabilitation: PET study of temporal voice activations. J. Nucl. Med. 49, 60–67. https://doi.org/10.2967/jnumed.107.045454.

Everhardt, M.K., Sarampalis, A., Coler, M., Rasker, D., Lowie, W., 2020. Meta-analysis on the identification of linguistic and emotional prosody in cochlear implant users and vocoder simulations. Ear Hear. 41, 1092–1102. https://doi.org/10.1097/ahu.0000000000001264.

Finke, M., Sandmann, P., Bonitz, H., Kral, A., Büchner, A., 2016. Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. Audiology and Neurotology 21, 305–315. https://doi.org/10.1159/000452123.

Friesen, L.M., Picton, T.W., 2010. A method for removing cochlear implant artifact. Hearing Res. 259, 95–106. https://doi.org/10.1016/j.heares.2009.10.012.

Giley, P.M., Sharma, A., Dorman, M., Finley, C.C., Panch, A.S., Martin, K., 2006. Minimization of cochlear implant stimulus artifact in cortical auditory evoked potentials. Clin. Neurophysiol. 117, 1772–1782. https://doi.org/10.1016/j.clinph.2006.04.018.
Wagner, L., Plontke, S.K., Rahne, T., 2017. Perception of iterated Rippled Noise Periodicity in Cochlear Implant Users. Audiology and Neurotology 22, 104–115. https://doi.org/10.1159/000478649.

Wedekind, A., Távora-Vieira, D., Nguyen, A.T., Marinovic, W., Rajan, G.P., 2021. Cochlear implants in single-sided deaf recipients: near normal higher-order processing. Clin. Neurophysiol. 132, 449–456. https://doi.org/10.1016/j.clinph.2020.11.038.

WHO, 2021. World report on hearing. https://www.who.int/publications/i/item/world-report-on-hearing.

Wilson, B.S., Dorman, M.F., 2008. Cochlear implants: a remarkable past and a brilliant future. Hearing Res. 242, 3–21. https://doi.org/10.1016/j.heares.2008.06.005.

Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited—again. Neuroimage 2, 173–181. https://doi.org/10.1006/nimg.1995.1025.

Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11, 946–953. https://doi.org/10.1093/cercor/11.10.946.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46. https://doi.org/10.1016/S1364-6613(01)01816-7.

Zhou, X., Seghouane, A.-K., Shah, A., Innes-Brown, H., Cross, W., Litovsky, R., McKay C. M., 2018. Cortical speech processing in postlingually deaf adult cochlear implant users, as revealed by functional near-infrared spectroscopy. Trends in Hearing. 22, 2331216518786850. https://doi.org/10.1177/2331216518786850.