Current Research in Neurobiology 3 (2022) 100059

Contents lists available at ScienceDirect

Current Research in Neurobiology

journal homepage: www.sciencedirect.com/journal/current-research-in-neurobiology

Electrophysiological differences and similarities in audiovisual speech processing in CI users with unilateral and bilateral hearing loss

Natalie Layer [a,*], Anna Weglage [a], Verena Müller [a], Hartmut Meister [b], Ruth Lang-Roth [a], Martin Walger [a,b], Micah M. Murray [c,d,e,f], Pascale Sandmann [a]

[a] University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Pediatric Audiology, Cochlear Implant Center, Germany
[b] Jean-Uhrmacher-Institute for Clinical ENT Research, University of Cologne, Germany
[c] The Sense Innovation and Research Center, Lausanne and Sion, Switzerland
[d] The LINE (The Laboratory for Investigative Neurophysiology), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[e] CIBM Center for Biomedical Imaging of Lausanne and Geneva, Lausanne, Switzerland
[f] Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Keywords:
Cochlear implant
Single-sided-deafness
Bilateral hearing loss
Event-related potential
Cortical plasticity
Multisensory integration
Audiovisual speech perception

ABSTRACT

Hearing with a cochlear implant (CI) is limited compared to natural hearing. Although CI users may develop compensatory strategies, it is currently unknown whether these extend from auditory to visual functions, and whether compensatory strategies vary between different CI user groups. To better understand the experience-dependent contributions to multisensory plasticity in audiovisual speech perception, the current event-related potential (ERP) study presented syllables in auditory, visual, and audiovisual conditions to CI users with unilateral or bilateral hearing loss, as well as to normal-hearing (NH) controls. Behavioural results revealed shorter audiovisual response times compared to unisensory conditions for all groups. Multisensory integration was confirmed by electrical neuroimaging, including topographic and ERP source analysis, showing a visual modulation of the auditory-cortex response at N1 and P2 latency. However, CI users with bilateral hearing loss showed a distinct pattern of N1 topography, indicating a stronger visual impact on auditory speech processing compared to CI users with unilateral hearing loss and NH listeners. Furthermore, both CI user groups showed a delayed auditory-cortex activation and an additional recruitment of the visual cortex, and a better lip-reading ability compared to NH listeners. In sum, these results extend previous findings by showing distinct multisensory processes not only between NH listeners and CI users in general, but even between CI users with unilateral and bilateral hearing loss. However, the comparably enhanced lip-reading ability and visual-cortex activation in both CI user groups suggest that these visual improvements are evident regardless of the hearing status of the contralateral ear.

1. Introduction

A cochlear implant (CI) can help restore the communication abilities in patients with severe to profound sensorineural hearing loss by electrically stimulating the auditory nerve (Zeng, 2011). However, listening with a CI is completely different from conventional hearing, as the electrical signal provided by the CI transmits only a limited amount of spectral and temporal information (Drennan and Rubinstein, 2008). Consequently, the central auditory system must learn to interpret the artificially sounding CI input as meaningful information (Giraud et al., 2001c; Sandmann et al., 2015). The ability of the nervous system to adapt to a new type of stimulus is an example of neural plasticity (Glennon et al., 2020; Merzenich et al., 2014). This phenomenon has been investigated in various studies with CI users, manifesting as an increase in activation in the auditory cortex to auditory stimuli during the first months after CI implantation (Giraud et al., 2001c; Green et al., 2005; Sandmann et al., 2015). Additional evidence for neural plasticity in CI users comes from the observation that these individuals recruit

* Corresponding author. at: University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Paediatric Audiology, Cochlear Implant Center, Kerpener Strasse 62, 50937, Cologne, Germany.
E-mail address: natalie.layer@uk-koeln.de (N. Layer).

https://doi.org/10.1016/j.crneur.2022.100059
Received 25 April 2022; Received in revised form 24 August 2022; Accepted 7 October 2022
Available online 8 November 2022

visual cortices for purely auditory speech tasks (Chen et al., 2016; Giraud et al., 2001c); a phenomenon referred to as cross-modal plasticity (e.g. Glennon et al., 2020).

Previous research has shown that event-related potentials (ERPs) derived from continuous electroencephalography (EEG) are an adequate method for studying cortical plasticity in CI users (Beynon et al., 2005; Finke et al., 2016a; Sandmann et al., 2009, 2015; Schierholz et al., 2015, 2017; Sharma et al., 2002; Viola et al., 2012; Layer et al., 2022). The primary benefit of analysing ERPs is the high temporal resolution, which allows for the tracking of individual cortical processing steps (Biasiucci et al., 2019; Michel and Murray, 2012). For instance, the auditory N1 (negative potential around 100 ms after stimulus onset) and the auditory P2 ERPs (positive potential around 200 ms after stimulus onset) are at least partly generated in the primary and secondary auditory cortices (Ahveninen et al., 2006; Bosnyak et al., 2004; Näätänen and Picton, 1987). Current models of auditory signal propagation recognise that there is an underlying anatomy exhibiting a semi-hierarchical and highly parallel organisation (e.g. Kaas and Hackett, 2000). In terms of auditory ERPs this would suggest that prominent components, such as the N1–P2 complex, include generators not only within primary auditory cortices, but within a distributed network along the superior temporal cortices as well as fronto-parietal structures and even visual cortices. Moreover, top-down effects, such as attention or expectation of incoming auditory events mediated by the frontal cortex, can influence these auditory processes (Dürschmid et al., 2019). The majority of previous ERP studies with CI users have used auditory stimuli to show that the N1 and P2 ERPs have a reduced amplitude and a prolonged latency in comparison to normal-hearing (NH) individuals. This observation suggests that CI users have difficulties in processing auditory stimuli (Beynon et al., 2005; Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009) and is consistent with previous behavioural results of impaired auditory discrimination ability in CI users (Sandmann et al., 2010, 2015; Finke et al., 2016a,b).

Although multisensory conditions more likely represent everyday situations, only a few ERP studies so far have been conducted with audiovisual stimuli. These studies primarily concentrated on rudimentary, non-linguistic audiovisual stimuli (sinusoidal tones and white discs) and showed a prolonged N1 response, and a greater *visual* modulation of the *auditory* N1 ERPs in CI users compared to NH listeners (Schierholz et al., 2015, 2017). Our previous study extended these results to more complex audiovisual syllables (Layer et al., 2022), using electrical neuroimaging (Michel et al., 2004, 2009; Michel and Murray, 2012) to perform topographic and ERP source analyses. Unlike traditional ERP data analysis, which is based on waveform morphology at specific electrode positions, electrical neuroimaging is reference-independent and takes into account the spatial characteristics and temporal dynamics of the global electric field to distinguish between the effects of response strength, latency, and distinct topographies (Murray et al., 2008; Michel et al., 2009). By using this topographic analysis approach, we previously showed a group-specific topographic pattern at N1 latency and an enhanced activation in the visual cortex at N1 latency for CI users when compared to NH listeners (Layer et al., 2022). These observations confirm a recent report about alterations in audiovisual processing and a multisensory benefit for CI users, if additional (congruent) visual information is provided (Radecke et al., 2022).

Based on previous ERP and behavioural results, one might conclude that multisensory processes, in particular integration of auditory and visual speech cues, remain intact in CI users despite the limited auditory signal provided by the CI. Nevertheless, it remains unclear whether the enhanced visual impact on auditory speech processing applies to all of the CI users, given that large inter-individual differences (e.g. with regards to the hearing threshold in the contralateral ear) have not been taken into account. Most of the aforementioned studies have included CI users with *bilateral hearing loss* (e.g. Finke et al., 2016a; Sandmann et al., 2015; Schierholz et al., 2015; Radecke et al., 2022; Layer et al., 2022), either provided with a CI on both ears (CI + CI on contralateral side) or on one ear (CI + hearing aid on contralateral side). These CI users will be referred to as CI-CHD users (CHD = 'contralateral hearing device') in the following. However, over the last years, the clinical margins for CI indication have been extended to *unilateral hearing loss*, enabling the implantation of single-sided deaf (SSD) patients (CI + NH on contralateral side; Arndt et al., 2011; Arndt et al., 2017; Buechner et al., 2010). This CI user group is particularly interesting, as the signal quality of the input is very different for the two ears, leading to maximally asymmetric auditory processing (Gordon et al., 2013; Kral et al., 2013). The variable hearing ability in the contralateral ear across different CI groups may at least partly account for the large variability in speech recognition ability observed in CI users (Lazard et al., 2012). To better understand the factors contributing to this variability and to extend previous findings on CI-CHD users and NH listeners (Layer et al., 2022), the current study systematically compared the timecourse of auditory and audiovisual speech processing as well as the lip-reading abilities between different groups of CI users, in particular CI-CHD users and CI-SSD users, and to NH listeners. The inclusion of the additional group of patients crucially extends our previous study because it not only evaluates the transferability our previous results to different patient groups but also provides deeper insights into the influence of individual factors - specifically the hearing ability of the second ear - on audiovisual speech processing in CI users. This is noteworthy because literature comparing CI-SSD to bimodal or bilateral CI users is scarce. However, the few existing studies reported differences in speech-in-noise performance (Williges et al., 2019) and in situations with multiple concurrent speakers (Bernstein et al., 2016) between CI-SSD users and bimodal or bilateral CI users, respectively. But, given this first evidence for purely auditory situations, we hypothesised that further differences would emerge for audiovisual stimulation, which has yet to be reported.

Given that CI-SSD users have an intact ear on the contralateral side, it is reasonable that this NH ear serves as the main communication channel despite the advantages given by the CI (Kitterick et al., 2015; Ludwig et al., 2021). Therefore, we hypothesised that CI-SSD users are less influenced by visual information, benefit less from audiovisual input and show poorer lip-reading skills than CI-CHD users. However, we expected a delay in cortical responses in the CI-SSD group, similar to the group of CI-CHD users, when compared to NH individuals, based on previous results from studies with purely auditory stimuli comparing the CI and the NH ear (Finke et al., 2016b; Bönitz et al., 2018; Weglage et al., 2022).

2. Material and methods

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the medical faculty of the University of Cologne (application number: 18 − 257). Prior to data collection, all participants gave written informed consent, and they were reimbursed.

2.1. Participants

In total, twelve post-lingually deafened CI-SSD patients were invited to participate in this study to extend the results from our previous study (Layer et al., 2022) by including an additional subgroup of CI-SSD users. Among these participants, one had to be excluded due to poor EEG data quality (high artefact load), resulting in a total of eleven CI-SSD patients (two right-implanted). Accordingly, we selected post-lingually deafened CI-CHD patients and NH listeners from our previous study (Layer et al., 2022; $n = 11$ each) such that they matched the CI-SSD patients as best as possible by gender, age, handedness, stimulated ear and years of education. The matched subset datasets from our previous study of CI-CHD users and NH listeners were reused for the current study and were extended by newly acquired data from an additional group of CI-SSD users. The CI-CHD users were implanted either unilaterally (n = 2; all left-implanted using a hearing aid on the contralateral ear) or bilaterally

($n = 9$; two right-implanted). All CI users had been using their device continuously for at least one year prior to the experiment. For the experiment, only the ear with a CI was stimulated and in the case of bilateral implantation, the 'better' ear (the ear showing the higher speech recognition scores in the Freiburg monosyllabic test) was used as stimulation side.

Thus, for final analyses, thirty-three volunteers were included, with eleven CI-SSD patients (7 female, mean age: 56.5 years ± 9.2 years, range: 39 – 70, 10 right-handed), eleven CI-CHD patients (7 female, mean age: 61.4 years ± 9.7 years, range: 39 – 75, 11 right-handed) and eleven NH listeners (7 female, mean age: 60.1 years ± 10.1 years, range: 34 – 70, 11 righthanded). Detailed information on the implant system and the demographic data are provided in Table 1. To check that cognitive abilities were age-appropriate, the DemTect Ear test battery was used (Brünecke et al., 2018). All participants scored within the normal, age-appropriate range (13 – 18 points). In addition, the German Freiburg monosyllabic speech test (Hahlbrock, 1970) with a sound intensity level of 65 dB SPL (see Table 4 for scores) was used to assess speech recognition abilities. To obtain a hearing threshold (HT) of the contralateral ear, we measured the aided HT in CI-CHD users in free-field and the unaided HT of the NH ear of CI-SSD users with headphones (see Table 1). All participants were native German speakers, had normal or corrected-to-normal vision (assessed by the Landolt test according to the DIN-norm; Wesemann et al., 2010) and none of the participants had a history of psychiatric disorder. Their handedness was assessed by the Edinburgh inventory (Oldfield, 1971).

2.2. Stimuli

The stimuli in this study were identical to those used in our previous study (Layer et al., 2022) and they were presented in three different conditions: visual-only (V), auditory-only (A) and audiovisual (AV). Additionally, there were trials with a black screen only ('nostim'), to which the participants were instructed to not react. The stimuli were delivered using the Presentation software (Neurobehavioral Systems, version 21.1) and a computer in combination with a duplicated monitor (69 inch). The stimuli consisted of the two syllables /ki/ and /ka/ which are included in the Oldenburg logatome speech corpus (OLLO; Wesker et al., 2005). They were cut from the available logatomes from one speaker (female speaker 1, V6 'normal spelling style', no dialect). These two syllables in particular differed in their phonetic distinctive features (vowel place and height of articulation) in the vowel contrast (/a/ vs. /i/; Micco et al., 1995). These German vowels are different in terms of central frequencies of the first (F1) and second formant (F2) representing the highest contrast between German vowels (e.g. Obleser et al., 2003), making them easily distinguishable for CI users. Importantly, as we presented visual-only syllables as well, the chosen syllables not only highly differ in terms of auditory (phoneme) realisation, but also in their visual articulatory (viseme) realisation. A viseme is the visual equivalent of the phoneme: a static image of a person articulating a phoneme (Dong et al., 2003). The editing of the syllables was done with Audacity (version 3.0.2) by cutting and adjusting them to the same duration of 400 ms. The syllables were normalised (adjusted to the maximal amplitude) in Adobe Audition CS6 (version 5.0.2).

To create a visual articulation of the auditory syllables, we used the

Table 1
Demographic information on the CI participants; HT (hearing threshold; average over 500 Hz, 1 kHz, 2 kHz, 4 kHz); Stim. = stimulated; HL = hearing loss; m = male; f = female.

ID	Group	Sex	Age	Handedness	Fitting	HT (dB HL; contralateral ear)	Stim. ear	Etiology	Age at onset of HL (years)	CI use of the stim. ear (months)	CI manufacturer
1	CI-CHD	m	61	right	bilateral	32	left	unknown	41	15	MedEl
2	CI-CHD	f	75	right	bilateral	31	left	hereditary	57	30	Advanced Bionics
3	CI-CHD	f	39	right	bilateral	26	right	otosclerosis	24	17	Advanced Bionics
4	CI-CHD	f	70	right	bilateral	37	left	unknown	37	56	MedEl
5	CI-CHD	f	70	right	bilateral	37	left	meningitis	69	20	MedEl
6	CI-CHD	m	59	right	bimodal	85	left	unknown	49	33	Advanced Bionics
7	CI-CHD	f	63	right	bilateral	36	left	meningitis	20	106	Advanced Bionics
8	CI-CHD	f	64	right	bilateral	29	left	whooping cough	9	78	Cochlear
9	CI-CHD	m	53	right	bilateral	36	left	unknown	30	235	Cochlear
10	CI-CHD	f	58	right	bimodal	41	left	unknown	49	18	Advanced Bionics
11	CI-CHD	m	56	right	bilateral	35	right	hereditary	19	63	MedEl
12	CI-SSD	f	64	right	SSD	10	left	unknown	49	30	Cochlear
13	CI-SSD	f	40	right	SSD	10	right	sudden hearing loss	34	77	MedEl
14	CI-SSD	m	43	right	SSD	12	left	bike accident	42	12	Cochlear
15	CI-SSD	m	54	right	SSD	17	left	unknown	52	28	MedEl
16	CI-SSD	f	49	right	SSD	23	right	otosclerosis	39	19	MedEl
17	CI-SSD	m	59	left	SSD	17	left	sudden hearing loss	49	54	Cochlear
18	CI-SSD	f	57	right	SSD	12	left	sudden hearing loss	20	53	Cochlear
19	CI-SSD	m	62	right	SSD	13	left	sudden hearing loss	47	63	Cochlear
20	CI-SSD	f	62	right	SSD	15	left	otitis media	14	65	MedEl
21	CI-SSD	f	68	right	SSD	25	left	unknown	60	50	Cochlear
22	CI-SSD	f	52	right	SSD	22	left	hereditary	20	12	MedEl

MASSY (Modular Audiovisual Speech SYnthesizer; Fagel and Clemens, 2004), which is a computer-based video animation of a talking head. This talking head has been previously validated for CI users (Meister et al., 2016; Schreitmüller et al., 2018) and is an adequate tool to generate audiovisual and visual speech stimuli (Massaro and Light, 2004). To generate articulatory movements matching the auditory speech sounds, one has to provide files that transform the previously transcribed sounds into a probabilistic pronunciation model providing the segmentation and the timing of every single phoneme. This can be done by means of the web-based tool MAUS (Munich Automatic Segmentation; Schiel, 1999). To obtain a video file of the MASSY output, the screen recorder Bandicam (version 4.1.6) was used in order to save the finished video files. Finally, the stimuli were edited in Pinnacle Studio 22 (version 22.3.0.377), making video files of each syllable in each condition: 1) Audiovisual (AV): articulatory movements with corresponding auditory syllables, 2) Auditory-only (A): black screen (video track turned off) combined with auditory syllables, 3) Visual-only (V): articulatory movements without auditory syllables (audio track turned off). Each trial started with a static face (500 ms) and was followed by the video, which lasted for 800 ms (20 ms initiation of articulatory movements + 400 ms auditory syllable + 380 ms completion of articulatory movements). For further analyses, we focused on the moving face (starting 500 ms post-stimulus onset/after the static face), as the responses to static faces comparing NH listeners and CI users have been investigated previously (Stropahl et al., 2015).

Supplementary video related to this article can be found at https://doi.org/10.1016/j.crneur.2022.100059

In general, all participants were assessed monaurally, meaning that in CI-SSD users the CI-ear was measured, in bimodal CI users the CI-ear was measured and in bilateral CI users the better CI-ear was measured. Regarding the *CI-SSD patients*, we followed the procedures of a previous study (Weglage et al., 2022) and positioned the processor inside an aqua case (Advanced Bionics; https://www.advancedbionics.com) to specifically stimulate the CI-ear without risking a cross-talking to the NH ear. Note that stimulation via a loudspeaker is not possible in this group, because an ear-plug is not enough to cover and mask the NH-ear, as it only reduces the intensity level by maximally 30 dB (Park et al., 2021). We refused the option of using noise to mask the NH ear, as this noise would have to be very loud to fully mask the information. This stimulation option would rather represent a speech-in-noise condition and would differ much more from the other two groups. An insert earphone (3M E-A-RTONE 3A) was put inside a hole of the aqua case where it was placed right over the microphone of the CI. A long coil cable was used to connect the processor with the implant. Regarding the *CI-CHD users*, the stimuli were presented via a loudspeaker (Audiometer-Box, type: LAB 501, Westra Electronic GmbH) which was placed in front of the participant. The hearing aid or the CI at the contralateral side was removed during the experiment and the ear was additionally covered with an ear-plug. For *NH participants*, the ear of the matched CI user was stimulated via an insert earphone (3M E-A-RTONE 3A), and the contralateral ear was closed with an ear-plug as well to avoid a cross-talking to the contralateral ear. The stimuli were calibrated to 65 dB SPL to ensure that the intensity level was equal for each stimulation technique. All participants rated the perceived loudness of the syllables with a seven-point loudness rating scale (as used in Sandmann et al., 2009, 2010), to ensure that the syllable intensity was perceived at a moderate level of 60− 70 dB (Allen et al., 1990). The stimuli (video files) are provided as supplementary material and can be downloaded.

2.3. Procedure

The procedure was identical to our previous study (see Layer et al., 2022). The additional CI-SSD users were seated comfortably in an electromagnetically shielded and dimly lit booth at a viewing distance of 175 cm to the screen. The participants were instructed to discriminate as fast and as accurately as possible between the syllables /ki/ and /ka/. The given response was registered using a mouse, with each of the two buttons assigned to one syllable. The sides were counterbalanced across the participants to prevent confounds caused by the used finger.

For each condition (AV, A, V, 'nostim'), 90 trials each were presented per syllable, resulting in a total number of 630 trials (90 repetitions x 3 conditions (AV, A, V) x 2 syllables (/ki/, /ka/) + 90 'nostim'-conditions). Each trial began with a 'nostim'-condition or a static face of the talking head (500 ms) followed by a visual-only, auditory-only or an audiovisual syllable. Afterwards, a fixation cross was shown until the participant pressed a button. In the case of a 'nostim'-trial, the participants were asked not to respond to. The trials were pseudo-randomised such that no trial of the same condition and syllable appeared twice in a row. The experiment lasted for 25 min excluding breaks, composed of five blocks of approximately 5 min each. A short break was given after each block. To ensure that the task was understood by the participants, we presented a short practice block consisting of five trials per condition before starting the recording. An illustration of the experimental paradigm can be found in Fig. 1A.

To obtain further behavioural measures apart from the ones registered in the EGG task (hit rates, response times), we asked the additionally measured CI-SSD users to rate the exertion of performing the task after each experimental block by using the 'Borg Rating of Perceived Exertion'-scale (Borg RPE-scale; Williams, 2017). Further, we measured the lip-reading abilities by means of a behavioural lip-reading test consisting of monosyllabic words from the German Freiburg test (Hahlbrock, 1970) which were visualised by various speakers and filmed (Stropahl et al., 2015). The participants were asked to watch the short videos of the muted monosyllabic word performances and to report which word was understood. This test was used in previous studies with CI patients (Stropahl et al., 2015; Stropahl and Debener, 2017), as well as in our previous study with CI-CHD users and NH listeners (Layer et al., 2022), whose scores we compared to CI-SSD users in the current study.

2.4. EEG recording

Similar to our previous study (Layer et al., 2022), the EEG data of the additionally measured CI-SSD users were continuously recorded by means of 64 AG/AgCl ActiCap slim electrodes using a BrainAmp system (BrainProducts, Gilching, Germany) and a customised electrode cap with an electrode layout (Easycap, Herrsching, Germany) according to the 10-10 system. To record an electrooculogram (EOG), two electrodes were placed below and beside the left eye (vertical and horizontal eye movements, respectively). The nose-tip was used as reference, and a midline electrode placed slightly anterior to Fz served as ground. Data recording was performed using a sampling rate of 1000 Hz. The online analog filter was set between 0.02 and 250 Hz. Electrode impedances were maintained below 10 k Ω during data acquisition.

2.5. Data analysis

The subset data taken from our previous study including the newly acquired data of the CI-SSD users were analysed in MATLAB 9.8.0.1323502 (R2020a; Mathworks, Natick, MA) and R (version 3.6.3; R Core Team (2020), Vienna, Austria). Topographic analyses were carried out in CARTOOL (version 3.91; Brunet et al., 2011). Source analyses were performed in Brainstorm (Tadel et al., 2011). The following R packages have been used: ggplot2 (version 2.3.3) for creating plots; dplyr (version 1.0.4), tidyverse (version 1.3.0) and tidyr (version 1.1.3) for data formatting; ggpubr (version 0.4.0) and rstatix (version 0.7.0) for statistical analyses.

2.5.1. Behavioural data

In a first step, we collapsed the syllables /ki/ and /ka/ for each condition (A, V, AV), as they did not show substantial differences between each other (mean RTs ± one standard deviation of the mean: CI-CHD: /ki/= 620 ms ± 88.3 ms, /ka/= 611 ms ± 88.0 ms; CI-SSD: /ki/=

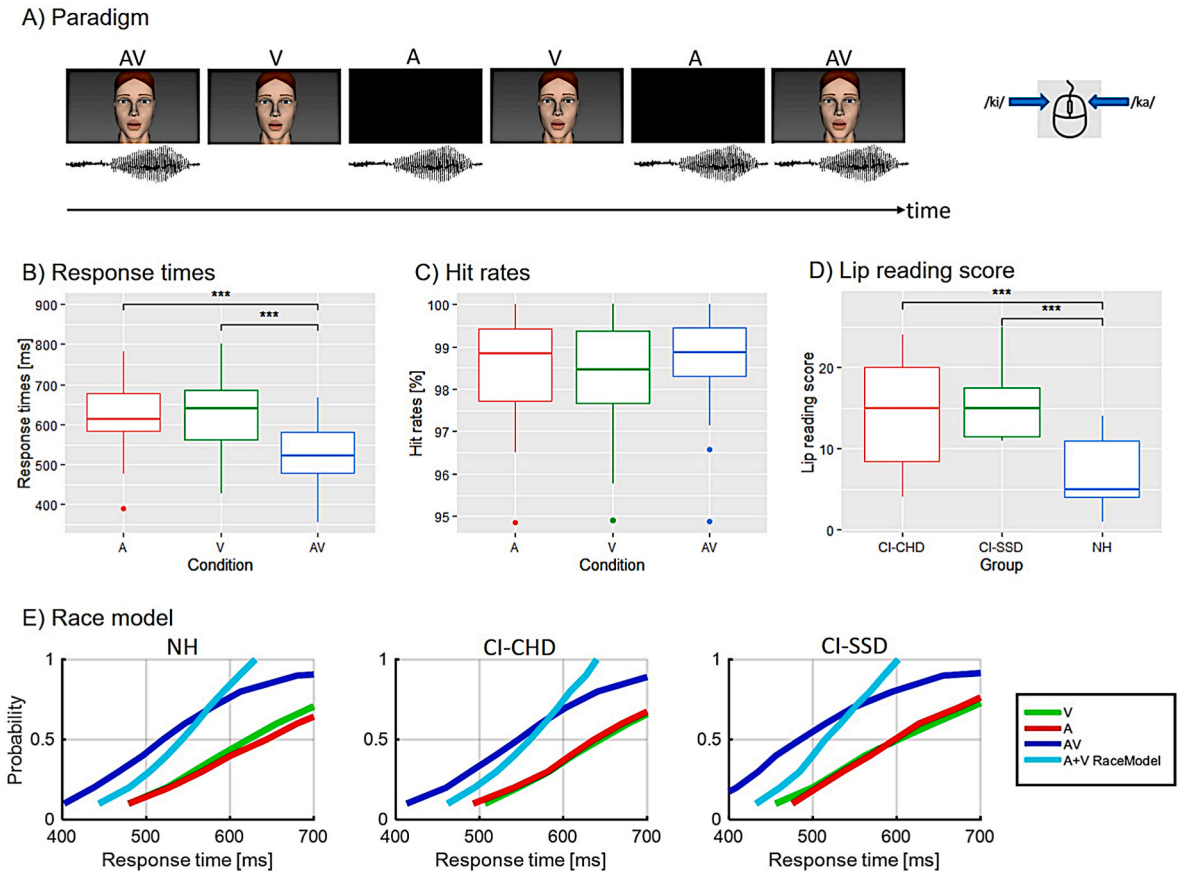

Fig. 1. Behavioural results. A) Overview of the paradigm (adapted from Layer et al., 2022). B) Mean response times for auditory (red), visual (green) and audiovisual (blue) syllables averaged over all groups, demonstrating that audiovisual syllables had shorter response times than auditory-only and visual-only RTs. C) Mean hit rates for auditory (red), visual (green) and audiovisual (blue) syllables averaged across all groups, with no differences between the three conditions. D) Cumulative distribution functions for CI-CHD, CI-SSD and NH. The race model is violated for all three groups because they show that the probability of faster response times is higher for audiovisual stimuli (blue line) than for those estimated by the race model (cyan line). Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

574 ms ± 100.0 ms, /ka/= 566 ms ± 94.0; NH: /ki/= 607 ms ± 91.6 ms, /ka/= 587 ms ± 102.0 ms; all $p \geq 0.354$). Second, false alarms or missing responses were discarded from the dataset. Trials that exceeded the individual mean by more than three standard deviations for each condition were declared as outliers and were removed from the dataset. Then, RTs and hit rates were computed for each condition (A, V, AV) for each individual. To analyse the performance for each condition and group, a 3 x 3 mixed ANOVA was used separately for the RTs and the hit rates, with condition (AV, A, V) as the within-subjects factor and group (NH, CI-CHD, CI-SSD) as the between-subjects factor. In the case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were carried out and corrected for multiple comparisons using a Bonferroni correction, in the case of significant main effects or interactions ($p \leq .05$). As the hit rates were very high in our previous study (Layer et al., 2022), we did not expect CI-SSD users to deviate from this pattern. Concerning the RTs, we expected similar results for CI-SSD users as for CI-CHD users and NH listeners, with shorter RTs for AV conditions compared to unisensory (A, V) conditions.

In a next step, we analysed the origin of the redundant signals effect, which is the effect of achieving faster RTs for audiovisual stimuli in comparison to unimodal stimuli (A, V) (Miller, 1982). For this purpose, we reused a subset of our previously reported data of the CI-CHD users and NH listeners (Layer et al., 2022) and extended these by the additional group of CI-SSD users. There are two accounts explaining this issue: the race model (Raab, 1962) and the coactivation model (Miller, 1982). Briefly, the race model claims that due to statistical facilitation it is more probable that either of the stimuli (A and V) will result in shorter response times in comparison to one stimulus alone (A or V). Therefore, one can assume that RTs of redundant signals (AV) are significantly faster, and that no neural integration is required to observe a redundant signals effect (Raab, 1962). In contrast, the coactivation model (Miller, 1982) assumes an interaction between the unimodal stimuli which forms a new product before initiating a motor response, leading to faster RTs. A widely used method in multisensory research is to test for the race model inequality (RMI; Miller, 1982) to explain whether the redundant signals effect was caused by multisensory processes or by statistical facilitation. According to the RMI, the cumulative distribution function (CDF) of the RTs in the multisensory condition (AV) can never exceed the sum of the CDFs of the two unisensory (A, V) conditions:

$$P(RT_{\mathrm{AV}} \leq t) \leq P(RT_{\mathrm{A}} \leq t) + P(RT_{\mathrm{V}} \leq t), \quad \text{for all } t \leq 0,$$

where $P(RT_x \leq t)$ represents the likelihood of a condition $x \in \{\mathrm{AV}, \mathrm{A}, \mathrm{V}\}$ being less than an arbitrary value t. Violation of this model, for any given value of t, is an indication for multisensory processes (see also Ulrich et al. (2007) for details). By applying the RMITest software by Ulrich et al. (2007), the CDFs of the RT distributions for each condition (AV, A, V) and for the sum of the modality-specific conditions (A + V) were estimated. The individual RTs were rank ordered for each condition to obtain percentile values (Ratcliff, 1979). Next, for each group separately (NH, CI-CHD, CI-SSD), the CDFs for the redundant signals conditions (AV) and the modality-specific sum (A + V) were compared for the five fastest deciles (bin width: 10 %). We used one-tailed paired t-tests

followed by a Bonferroni correction to account for multiple comparisons. Significance at any decile bin was treated as violation of the race model, suggesting multisensory interactions in the behavioural responses. Here, we expected a similar redundant signals effect for CI-SSD users, as CI-CHD users and NH listeners both showed a violation of the race model inequality in our previous study (Layer et al., 2022).

To assess differences between the CI user groups and the NH listeners in the lip-reading task and in the subjective rating of exertion, we performed one-way ANOVAs. Whenever a significant main effect of group was present, we performed follow-up tests with a Bonferroni correction to account for multiple comparisons. Concerning the lip-reading ability, we anticipated that CI-SSD users performed worse compared to CI-CHD users due to their intact contralateral ear, which may reduce the need to rely on lip movements in their everyday life. In terms of subjective exertion rating, we expected no difference between experimental groups because our previous study (Layer et al., 2022) found no difference, which was likely due to the easy task.

2.5.2. EEG pre-processing

The pre-processing of the EEG data was done with EEGLAB (version v2019.1; Delorme and Makeig, 2004), a software working within the MATLAB environment (Mathwork, Natick, MA). The raw data were down-sampled (500 Hz) and filtered with a FIR-filter, having a high pass cut-off frequency of 0.5 Hz with a maximum possible transition bandwidth of 1 Hz (cut-off frequency multiplied by two), and a low pass cut-off frequency of 40 Hz with a transition bandwidth of 2 Hz. For both filters, the Kaiser-window (Kaiser-$\beta = 5.653$, max. stopband attenuation $= -60$ dB, max. passband deviation $= 0.001$; Widmann et al., 2015) was used to maximise the energy concentration in the main lobe by averaging out noise in the spectrum and minimising information loss at the edges of the window. Electrodes in the proximity of the speech processor and transmitter coil were removed for CI users (mean: 2.8 electrodes; range: $1 - 4$). Afterwards, the datasets were epoched into 2 s dummy epoch segments, and pruned of unique, non-stereotype artefacts using an amplitude threshold criterion of four standard deviations. An independent component analysis (ICA) was computed (Bell and Sejnowski, 1995) and the resulting ICA weights were applied to the epoched original data ($1 - 40$ Hz, -200 to 1220 ms relative to the stimulus onset (including the static and moving face)). Independent components reflecting vertical and horizontal ocular movements, electrical heartbeat activity, as well as other sources of non-cerebral activity were rejected (Jung et al., 2000). Independent components exhibiting artefacts of the CI were identified based on the side of stimulation and the time course of the component activity, showing a pedestal artefact around 700 ms after the auditory stimulus onset (520 ms). The identified components were removed from the EEG data. In a next step, we interpolated the missing channels using a spherical spline interpolation (Perrin et al., 1989) which allows for a solid dipole source localisation of auditory ERPs in CI users (Debener et al., 2008; Sandmann et al., 2009). Only trials yielding correct responses (NH: 91.0 % $\pm$ 3.9 %; CI-CHD: 88.1 % $\pm$ 3.8 %; CI-SSD: 86.6 % $\pm$ 4.5 %) were kept for further ERP analyses.

2.5.3. EEG data analysis

We compared event-related potentials (ERPs) of all conditions (AV, A, V) between the two CI user groups and the NH participants. The additive model which is denoted by the equation $AV = A + V$ (Barth et al., 1995) was used to investigate multisensory interactions. The model is satisfied and suggests independent and linear processing if the multisensory (AV) responses equal the sum of the unisensory (A, V) responses. Whereas, if the model is not satisfied, non-linear interactions between the unisensory modalities are assumed (Barth et al., 1995). Similar to our previous study (Layer et al., 2022), we rearranged the equation to $A = AV - V$ such that we could compare the directly measured auditory ERP response (A) with the term [AV $-$ V], denoting an ERP difference wave representing a visually-modulated auditory ERP response. Hence, [AV-V] is an estimate of an auditory response elicited in a multisensory context. In the case of a lack of interaction between the two unisensory (A, V) modalities, both A and AV-V should be identical. However, if the auditory (A) and the modulated auditory (AV-V) ERPs are *not* identical, this would point to non-linear multisensory interactions (Besle et al., 2004; Murray et al., 2005; Cappe et al., 2010; Foxe et al., 2000; Molholm et al., 2002). Such non-linear effects can be either super-additive ($A < AV - V$) or sub-additive ($A > AV - V$). But, since interpreting these effects is not straightforward, it is necessary to obtain reference-independent measurements of power or of source estimates (e.g. Cappe et al., 2010). Before creating the difference waves (AV-V), we randomly reduced the number of epochs based on the condition with the lowest number of epochs for each individual to guarantee that there was an equal contribution of each condition to the resulting difference wave. The difference waves were only created for the CI-SSD users in this study, and the difference waves for the NH listeners and CI-CHD users were reused from our previous study (Layer et al., 2022).

As in our previous study (Layer et al., 2022), we analysed our ERP data within an electrical neuroimaging framework (Murray et al., 2008; Michel et al., 2009; Michel and Murray, 2012), comprising topographic and ERP source analysis to compare auditory (A) and modulated (AV-V) ERPs within and between groups (NH, CI-CHD, CI-SSD). We investigated the global field power (GFP) and the global map dissimilarity (GMD) to quantify ERP differences in response strength and response topography, respectively (Murray et al., 2008). First, we looked at the GFP, at the time window of the N1 and the P2 (N1: $80 - 200$ ms; P2: $200 - 370$ ms), which were chosen based on visual inspection of the GFP computed for the grand average ERPs across conditions and groups. The GFP is the spatial standard deviation of all electrode values at a specific time point (Murray et al., 2008) and was first described by Lehmann and Skrandies (1980). The reason for choosing the GFP instead of selecting specific channels of interest is that this approach allows for a more objective peak detection. The GFP peak mean amplitudes and latencies were detected for each individual, condition (A, AV-V) and time window (N1, P2) and were statistically analysed by using a 3 x 2 mixed ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) as the within-subjects factor for each peak separately. Based on previous observations with CI-CHD users (Beynon et al., 2005; Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009; Layer et al., 2022) and CI-SSD users (Finke et al., 2016b; Bönitz et al., 2018; Weglage et al., 2022), we expected delayed N1 and reduced P2 responses for all CI user groups compared to NH controls.

Second, we analysed the GMD (Lehmann and Skrandies, 1980) to quantify topographic dissimilarities (and by extension, dissimilar configurations of neural sources; Vaughan Jr, 1982) between experimental conditions and groups, regardless of the signal strength (Murray et al., 2008). The GMD was analysed in CARTOOL by computing a 'topographic ANOVA' (TANOVA; Murray et al., 2008) to quantify differences in topographies between groups for each condition. Even though the name is misleading, this is no analysis of variance, but a non-parametric randomisation test. This randomisation test was executed with 5.000 permutations and by calculating sample-by-sample p-values. To control for multiple comparisons, an FDR correction was applied (FDR = false discovery rate; Benjamini and Hochberg, 1995). Since ERP topographies remain stable for a certain period of time before changing to another topography (called 'microstates'; Michel and Koenig, 2018) and to account for temporal autocorrelation, the minimal significant duration was adjusted to 15 consecutive time frames, corresponding to 30 ms.

2.5.4. Hierarchical clustering and single-subject fitting analysis

Whenever differences in topographies (GMD) for two groups or conditions are found, this is an indication for distinct neural generators contributing to these topographies (e.g. Vaughan Jr, 1982). However, it is also possible that a GMD is caused by a latency shift of the ERP, meaning that the same topographies are present but just shifted in time (Murray et al., 2008). To disentangle these two possible GMD causes, we performed a hierarchical topographic clustering analysis with

group-averaged data (NH(A), NH(AV-A), CI-CHD(A), CI-CHD(AV-V), CI-SSD(A), CI-SSD(AV-V)) to identify template topographies within the time windows of interest (N1, P2). Again, we used CARTOOL for this analysis and chose the atomize and agglomerate hierarchical clustering (AAHC) which has been especially designed for ERP-data (detailed in Murray et al., 2008). This method includes the global explained variance of a cluster and prevents blind combinations (agglomerations) of short-duration clusters. Thus, this clustering method identifies the minimal number of topographies accounting for the greatest variance within a dataset (here NH(A), NH(AV-A), CI-CHD(A), CI-CHD(AV-V), CI-SSD(A), CI-SSD(AV-V)).

In a next step, the template maps detected by the AAHC were entered into a single-subject fitting (Murray et al., 2008) to identify the distribution of specific templates by calculating sample-wise spatial correlations for each individual and condition between each template topography and the observed voltage topographies. Each sample was matched to the template map with the largest spatial correlation. For statistical analyses, we focused on the *first onset of maps* (latency) and the *map presence* (number of samples in time frames) which are two of many other output options provided by CARTOOL. Particularly, we performed a mixed ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as within-subject factors, separately for each time window (N1, P2). In the case of significant three-way interactions, group-wise mixed ANOVAs (condition x template map) were computed. A Greenhouse-Geisser correction was applied whenever there was a violation of the sphericity assumption. Post-hoc t-tests were computed and corrected for multiple comparisons using a Bonferroni correction. We anticipated that the analysis of the first onset of maps would confirm a delayed N1 latency for both CI-CHD and CI-SSD users based on previous results. In terms of map presence at N1 latency range, we speculated that there would be a pattern between CI-CHD users and NH listeners for CI-SSD users, as they have both a CI and a NH ear. However, we are not aware of previous studies reporting similar results for CI-SSD users.

2.5.5. Source analysis

We performed an ERP source analysis for each group and condition by means of Brainstorm (Tadel et al., 2011) to find out whether topographic differences can be explained by fundamentally different configurations of neural generators. The tutorial provided by Stropahl et al. (2018) served as guideline for conducting the source analysis. As in our previous study (Layer et al., 2022) and in various studies with CI patients (Bottari et al., 2020; Stropahl et al., 2015; Stropahl and Debener, 2017), we selected the method of dynamic statistical parametric mapping (dSPM, Dale et al. (2000)). dSPM works more precisely in identifying deeper sources than standard norm methods, even though the spatial resolution stays relatively low (Lin et al., 2006). It takes the minimum-norm inverse maps with constrained dipole orientations to approximate the locations of electrical activity recorded on the scalp. This method can be successfully used to localise small cortical areas such as the auditory cortex (Stropahl et al., 2018). First, individual noise covariances were calculated from single-trial pre-stimulus onset baseline intervals (−50 to 0 ms) to estimate single-subject based noise standard deviations at each location (Hansen et al., 2010). As a head model, the boundary element method (BEM) which is implemented in OpenMEEG was used. The BEM gives three realistic layers and representative anatomical information (Gramfort et al., 2010). The final activity data is then displayed as absolute values with arbitrary units based on the normalisation within the dSPM algorithm. Consistent with the procedures of our previous study (Layer et al., 2022), we defined an *auditory* and a *visual* ROI by combining smaller regions within the Destrieux-atlas (Destrieux et al., 2010; Tadel et al., 2011; auditory: G_temp_sup-G_T_transv, S_temporal_transverse, G_temp_sup-Plan_tempo and Lat_Fis-post; visual: G_cuneus, S_calcarine, S_parieto_occipital). These ROIs were chosen in accordance with several previous studies (Stropahl et al., 2015; Stropahl and Debener, 2017; Giraud et al., 2001b, c; ; Prince et al., 2021; Layer et al., 2022). In specific, the chosen parts of the auditory ROI have been reported as both N1 (Näätänen and Picton, 1987; Godey et al., 2001; Woods et al., 1993; Bosnyak et al., 2004) and P2 (Crowley and Colrain, 2004 (for review); Hari et al., 1987; Bosnyak et al., 2004; Ross and Tremblay, 2009) generators. The selected ROIs can be viewed in Fig. 4A.

Source activities for each ROI, condition and group were exported from Brainstorm for each participant. Afterwards, the peak means and latencies for each time window of interest (N1: 80 − 200 ms, P2: 200 − 370 ms) were extracted. A mixed-model ANOVA was performed separately for each time window with group (NH, CI-CHD, CI-SSD) as between-subject factor and condition (A, AV-V), ROI (auditory, visual) and hemisphere (left, right) as within-subjects factors. A Greenhouse-Geisser correction was applied in the case of violation of the sphericity assumption. In the case of significant interactions or main effects, post-hoc t-tests were computed and corrected for multiple comparisons using a Bonferroni correction. Similar to our hypotheses for the fitting data, we speculated that a pattern between the one of CI-CHD users and NH listeners would emerge for the recruitment of the visual cortex, which we observed for CI-CHD users in our previous study (Layer et al., 2022). In addition, in accordance with the fitting data and the GFP, we expected a delayed auditory cortex response for CI-SSD users as well. Finally, based on our previous study, we expected to find indications for multisensory processing, with different activity for AV-V compared to A for CI-SSD users, too.

3. Results

3.1. Behavioural results

Overall, all participants showed hit rates of ≥ 98 % in all conditions, and the mean RTs were between 504 ms and 638 ms (see Table 2). The 3 x 3 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI-CHD, CI-SSD, NH) as between-subject factor revealed for RTs no main effect of group ($F_{2,30} = 1.06$; $p = .36$, $\eta p^2 = 0.058$) and no group × condition interaction, but a main effect of condition ($F_{1.31,39.3} = 100.63$; $p \leq .001$, $\eta p^2 = 0.31$). Follow-up post-hoc t-tests showed that RTs to redundant signals (AV) were significantly faster when compared to V ($t(32) = 12.4$; $p \leq .001$) or A ($t(32) = 20.3$; $p \leq .001$). There was no difference in RTs between the unisensory stimuli A and V ($t(32) = -0.84$; $p = 0.41$). These results are displayed in Fig. 1B.

For the hit rates, the 3 x 3 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI-CHD, CI-SSD, NH) as between-subject factor showed no main effects or interactions (see Fig. 1C).

Concerning the race model, the one sample t-tests were significant in at least one decile for each group (see Table 3). This means that the likelihood of faster response times for redundant signals (AV) is higher than for those estimated by the race model (A + V). Fig. 1D displays the results of the race model. Overall, the violation of the race model in CI-CHD, CI-SSD users and NH listeners confirms the existence of multisensory integration in all tested groups.

For the other behavioural measures, we calculated one-way ANOVAs with subsequent t-tests to assess differences in *auditory word recognition ability* and *(visual) lip-reading abilities* between CI-CHD, CI-SSD users and

Table 2
Mean hit rates (in %) and mean response times (in ms).

Condition	Hit rates			Response times		
	NH	CI-CHD	CI-SSD	NH	CI-CHD	CI-SSD
A	99.0 ± 0.7	98.6 ± 1.1	98.1 ± 1.6	638 ± 84.1	623 ± 92.4	607 ± 79.4
V	98.5 ± 1.2	98.0 ± 1.3	98.5 ± 1.3	638 ± 98.4	623 ± 96.1	602 ± 87.7
AV	98.7 ± 1.0	98.3 ± 1.3	99.1 ± 0.8	526 ± 81.9	530 ± 89.4	504 ± 74.0

Table 3
Redundant signals and modality-specific sum in each decile. AV is the redundant signals condition. A + V is the modality-specific sum. Paired-samples one-tailed t-tests were conducted for each group (with Bonferroni correction for multiple comparisons). An asterisk indicates a statistically significant result ($p \leq .05/5$).

Decile	NH			CI-CHD			CI-SSD		
	AV	A + V	p	AV	A + V	p	AV	A + V	p
.10	402	444	.000*	413	462	.000*	373	432	.000*
.20	437	480	.000*	460	494	.008*	409	461	.000*
.30	467	504	.000*	489	520	.014	435	484	.000*
.40	496	525	.003*	519	541	.056	456	450	.000*
.50	520	543	.011	547	560	.169	485	515	.000*

Table 4
Other behavioural measures for CI users and NH listeners. In the Freiburg monosyllabic word test and in the lip-reading test, a score of 100% means that all words have been repeated correctly. A higher value for the exertion rating means it was more effortful to perform the task (range: 6–20; 6 = no effort, 20 = highly effortful).

Group	Freiburg test (%)	Lip-reading test	Exertion rating
CI-CHD	70.9 ± 12.2	14.5 ± 6.7	12.5 ± 2.0
CI-SSD	71.4 ± 13.1	15.5 ± 4.4	12.2 ± 1.8
NH	98.2 ± 4.0	6.9 ± 4.7	11.9 ± 1.6

NH listeners. The ANOVAs showed a main effect of group for both the auditory word recognition ability ($F_{2,30} = 23.99$; $p < .001$, $\eta p^2 = 0.615$) and (visual) lip-reading abilities ($F_{2,30} = 8.6$; $p = .001$, $\eta p^2 = 0.364$). Follow-up t-tests revealed poorer speech recognition ability in the Freiburg monosyllabic test ($p \leq .001$), but better lip-reading skills for all CI users when compared to NH listeners (CI-CHD vs. NH: $p = .002$, CI-SSD vs. NH: $p = .004$). There was no difference between the two CI-user groups (CI-CHD vs. CI-SSD: $p = 0.6$). Concerning the *subjective exertion* measured during the EEG task, the ANOVA did not show a difference between the CI-CHD, the CI-SSD users and NH listeners ($F_{2,30} = 0.25$; $p = 0.78, \eta p^2 = 0.016$). These results indicate that none of the tested groups perceived the task as more effortful than another group. The scores of these tests can be found in Table 4.

3.2. ERP results on the sensor level: GFP

In Fig. 2 the GFP of the grand averaged ERPs for the unisensory auditory (A) and the visually modulated auditory (AV-V) responses are shown for each group. Approximately between 120 and 140 ms, the first prominent peak is visible for all three groups. This peak fits into the time window of a N1 ERP. The next peak is around 240 ms and seems to be more prominent in NH listeners when compared to the two CI user groups. In the following, this peak is labelled as the P2 ERP. The GFP of the other conditions (V, AV) are also shown in the supplementary material (including the GMD between groups for each condition) to give an idea of the "raw," non-difference wave data. First, we calculated a 3 x 2 mixed ANOVA with group (NH, CI-CHD, CI-SSD) as between-subject variable and condition (A, AV-V) as within-subject factor for the N1 GFP peak mean amplitude and the GFP peak latency. For the *N1 peak amplitude*, no statistically significant main effects or interactions were found. However, the ANOVA with *N1 latency* revealed a significant main effect of group ($F_{2,30} = 4.76$; $p \leq .05$, $\eta p^2 = 0.18$). Follow-up t-tests

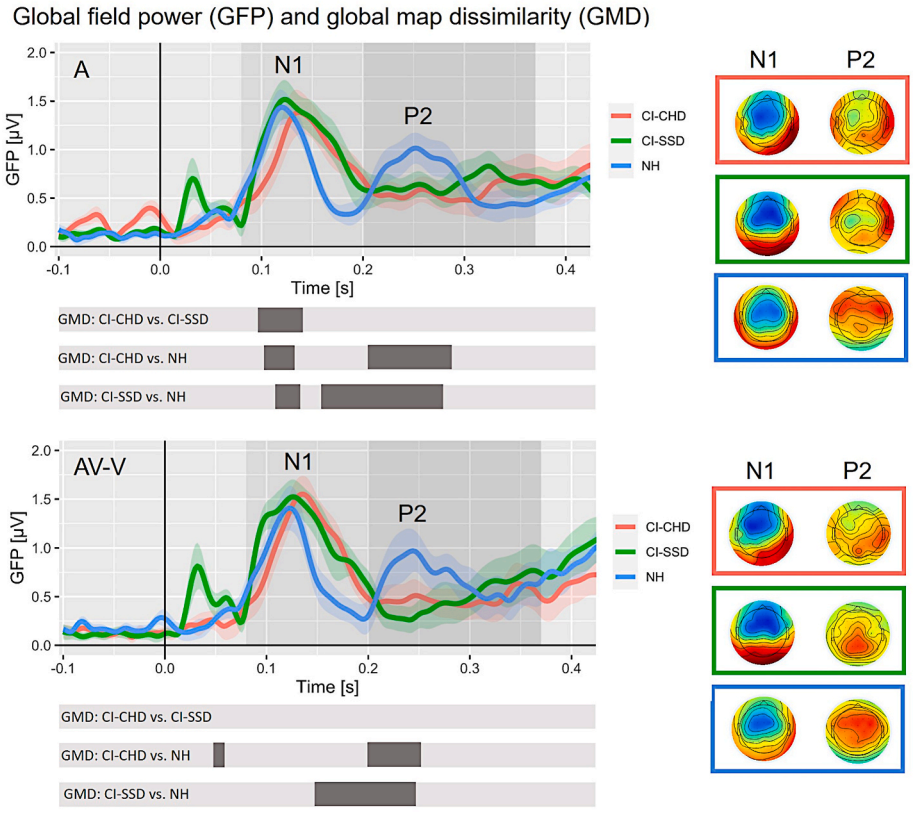

Fig. 2. ERP results on the sensor level. A) GFP of conditions A and AV-V for CI-CHD users (red), CI-SSD users (green) and NH listeners (blue), including standard error. It is important to note that the GFP only provides positive values because it represents the standard deviation across all electrodes separately for each time point. The ERP topographies at the GFP peaks (N1(A) = CI-CHD: 147 ms, CI-SSD: 136 ms, NH: 118 ms; N1(AV-V) = CI-CHD: 137 ms, CI-SSD: 135 ms, NH: 118 ms; P2(A) = CI-CHD: 305 ms, CI-SSD: 288 ms, NH: 256 ms; P2(AV-V) = CI-CHD: 284 ms, CI-SSD: 307 ms, NH: 245 ms) are given separately for each group (displayed on the right). The grey-shaded areas represent the N1 and P2 time windows for detecting peak and latency. The grey bars below represent the time window in which significant GMDs between the three groups were observed. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

showed a prolonged N1 latency for both the CI-CHD ($t(21) = 3.14$; $p \leq .05$) and the CI-SSD users ($t(21) = 3.83$; $p \leq .05$) compared to NH individuals. There was no significant difference between the two CI groups ($t(21) = -0.03$; $p = .97$).

We performed the same 3 x 2 mixed ANOVA for the P2 GFP peak mean amplitude and latency. For both the P2 peak amplitude and latency, there was no significant main effects or interactions.

3.3. ERP results on the sensor level: GMD

The GMD was analysed sample-by-sample to identify if and when ERP topographies significantly differ between conditions and groups. We compared CI-CHD with NH listeners (CI-CHD vs. NH), CI-SSD with NH listeners (CI-SSD vs. NH) and both CI groups (CI-CHD vs. CI-SSD) *separately for each condition (A and AV-V)*. For the *auditory condition (A)*, the results revealed topographic differences for all group comparisons within the time window of the N1 (CI-CHD vs. NH: 90− 132 ms, CI-SSD vs. NH: 114 – 128 ms, CI-CHD vs. CI-SSD: 100 – 122 ms). Concerning the topographic differences within the P2 time window, there were no differences between the two groups of CI users, but we observed differences between the NH listeners and the two CI groups (CI-CHD vs. NH: 208 – 280 ms, CI-SSD vs. NH: 152 – 272 ms).

Regarding the GMD for the *modulated condition (AV-V)*, there was a difference between NH listeners and CI-CHD at the N1 time window (NH vs. CI-CHD: 104 – 126 ms). Within the P2 time window, again there were no differences between the two CI groups, however there were differences between the NH listeners and each CI group (CI-CHD vs. NH: 204 – 264 ms, CI-SSD vs. NH: 154 – 242 ms). In addition, the GMD duration at the P2 time window was shorter for AV-V compared to A. The exact durations displaying differences between the groups are illustrated in Fig. 2 (grey bars beneath the GFP plots).

3.4. ERP results on the sensor level: Hierarchical clustering and single-subject fitting results

To better understand the underlying topographic differences (i.e. GMD) between the three groups, we conducted a hierarchical topographic clustering analysis by using the group-averaged data (CI-CHD (A); CI-CHD(AV-V); CI-SSD(A); CI-SSD(AV-V); NH(A); NH(AV-V)) in order to find template topographies within the N1 and P2 time windows. For that purpose, we chose a segment ranging from −100 ms to 470 ms (50–235 time frames). Specifically, we employed the atomize and agglomerate hierarchical clustering (AAHC) to identify the minimal amount of topographies that can explain the variance in our data set as best as possible. This method detected 17 template maps in 18 clusters that explained 88.08 % of all data. To be precise, we detected two maps within the N1 time window (map A and Map B) and three prominent maps within the P2 time window (Map C, Map D, Map E). With these template maps, we performed a single-subject analysis (Murray et al., 2008) to determine how well each of the template maps spatially correlated with the data from each participant. As the *template Map B* matches the topography from a conventional N1 peak (Fig. 2; Finke et al., 2016a; Sandmann et al., 2015), this template map will be referred to as N1 topography hereafter. *Template Map C* looks like a typical P2 topography (Fig. 2; Finke et al., 2016a; Schierholz et al., 2021) and therefore we will refer to this template map as the P2 topography. The *template Map E* is particularly prominent in the two CI-user groups and will be referred to as P2-like topography due to its similarity to the P2 topography (Fig. 2).

Dissimilarities within the topography across groups and conditions (see section 'ERP results on the sensor level: GMD') can be explained by a latency shift of the ERPs and/or by distinct neural generator configurations. To shed light on the origin of these differences, we analysed the *first onset of maps* and the *map presence* for the N1 and the P2 time windows. These results are presented in the following two subsections.

3.4.1. N1 time window

On the *descriptive level*, the CI-CHD users showed a map A and a Map B (= N1 topography) which were both present in the auditory-only condition (A; number of samples map A: 18.3 ± 18.6; Map B: 30.5 ± 19.0). Interestingly, specifically in the modulated condition (AV-V), the Map B (= N1 topography) was clearly more frequent compared to map A (number of samples 39.9 ± 18.2 (Map B) vs. 9.27 ± 16.6 (map A)). By contrast, both the NH listeners and the CI-SSD users showed a greater presence of Map B (= N1 topography) in general, irrespective of condition (A: number of samples: NH 40.8 ± 10.8; CI-SSD 48.0 ± 6.71; AV-V: number of samples: NH 36.9 ± 12.5; CI-SSD 47.5 ± 8.26).

To obtain an explanation for a potential ERP latency shift, we statistically analysed the *first onset of maps* by using a mixed-model ANOVA with group (NH, CI-CHD; CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subject factors for the N1 time window. The three-way mixed ANOVA revealed a significant group × map interaction ($F_{2,29} = 13.4$; $p \leq .001$, $\eta p^2 = 0.39$). Follow-up t-tests showed that the onset of Map B (=N1 topography) was earlier in the NH listeners when compared with the CI-CHD ($t(20) = 3.82$; $p \leq .01$) and the CI-SSD ($t(21) = 6.75$; $p \leq .001$). There was no group difference in the onset of the N1 topography between CI-CHD and CI-SSD ($t(20) = -0.54$; $p = .6$). The results suggest that the N1 is generated later in CI users compared to NH individuals, regardless of the hearing threshold of the contralateral ear.

Second, we statistically analysed the number of time frames of the maps that showed the highest spatial correlations to the single-subject data, i.e. the *map presence*. This variable can provide an explanation for potentially distinct underlying neural generators between the three groups (CI-CHD, CI-SSD, NH) and the two conditions (A, AV-V). As above, we calculated a mixed-model ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subjects factors for the N1 time window.

For the N1 template maps, the ANOVA results showed a group x map × condition interaction ($F_{2,30} = 7.12$; $p \leq .005$, $\eta p^2 = 0.04$). Post-hoc t-tests revealed for the CI-CHD users, but not for the CI-SSD users or NH listeners, that the presence of Map B (=N1 topography) was significantly *enhanced* for the modulated (AV-V) compared to the auditory-only (A) condition ($t(10) = -2.85$; $p \leq .05$). These results are illustrated in Fig. 3A. Given that template Map B corresponds to a conventional N1 topography, the results suggest that CI-CHD users in specific generate a N1 ERP map for the modulated response (AV-V) more frequently compared to the unisensory (A) condition. This visual modulation effect at the N1 latency was not observable for the NH listeners and the CI-SSD users.

Taken together, our results for the N1 on the *first onset of maps* and the *map presence* suggest that the observed topographic group differences at N1 latency can be explained by the following two reasons: 1) there are generally delayed cortical N1 ERPs in CI users, regardless of the condition (auditory-only or modulated response) and regardless of whether these patients have unilateral or bilateral hearing loss, and 2) there is a distinct pattern of ERP topographies specifically for the CI-CHD users compared to NH listeners and CI-SSD users. The visual modulation effect in the N1 topography was only observed for CI-CHD users, which suggests that this CI group in particular has a strong visual impact on auditory speech processing. By contrast, the visual impact in the CI-SSD users seems to be less pronounced and appears to be comparable to the NH listeners.

3.4.2. P2 time window

Similar to the analysis on the N1 time window, we analysed the *first onset of maps* by using a mixed-model ANOVA with group (NH, CI-CHD; CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subject factors for the P2 time window. The results did not reveal any significant main effects or interactions.

In a second step, we analysed the *map presence* for the P2 time window. As above, we calculated a mixed-model ANOVA with group

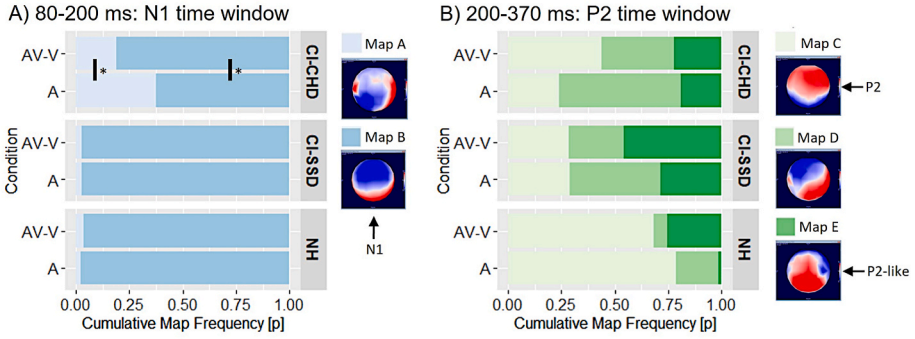

Fig. 3. Results from the hierarchical clustering and the single-subject fitting. A) Cumulative map frequency of the N1 maps: the CI-CHD users, but not the NH listeners or CI-SSD users, show a condition effect, with more frequent N1 map presence for AV-V compared to A. The corresponding map topographies are displayed on the right side, with Map B being referred to as the N1 topography. B) Cumulative map frequency of the P2 maps: there is a group effects (independent of the condition): NH listeners reveal a more frequent presence of a P2 topography (Map C) compared to CI-SSD users, and CI-SSD users show a more frequent presence of a P2-like topography (Map E) compared to NH listeners. CI-CHD users show a more frequent presence of Map D compared to NH listeners. Additionally, there is a condition effect (independent of the group): The presence of the P2-like topography (Map E) is enhanced for AV-V compared to A. This suggests a visual modulation of auditory speech processing at P2 latency in all groups. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

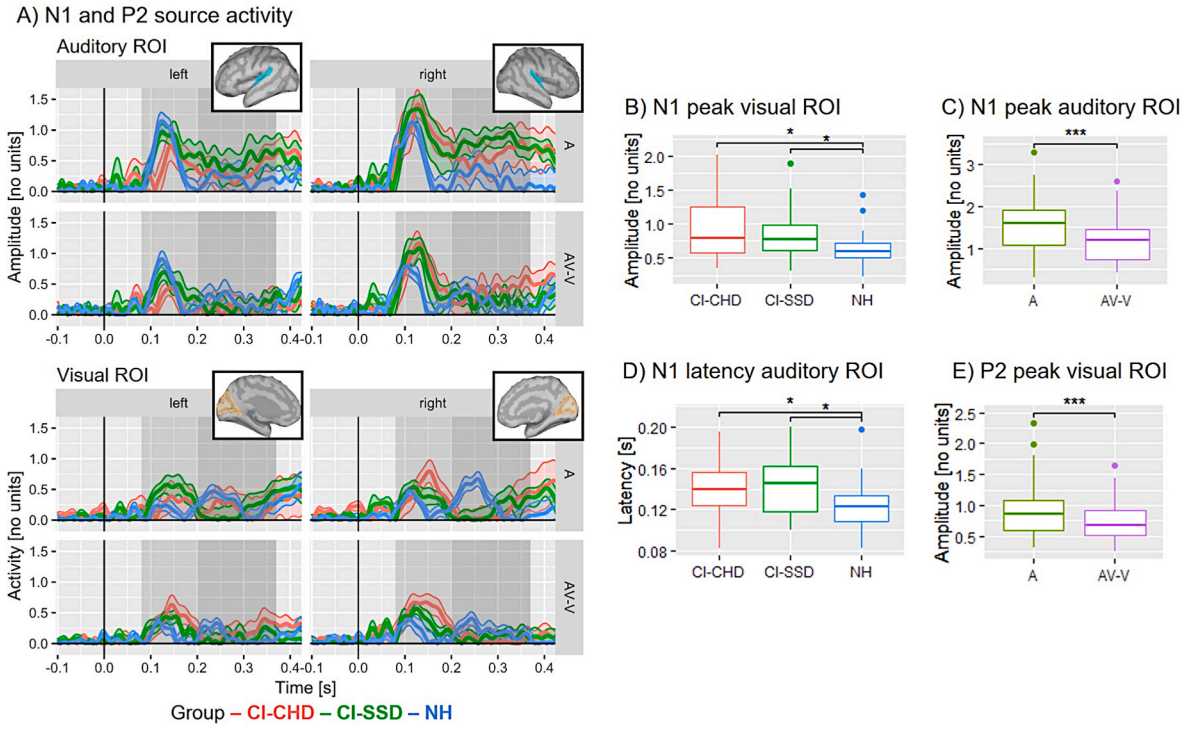

Fig. 4. ERP results on the source level. A) N1 and P2 source activity for CI users (red), CI-SSD users (green) and NH listeners (blue) separately for each ROI and each hemisphere with standard error (standard error shading was capped at zero). The source activity is displayed as absolute values with arbitrary units based on the normalisation within Brainstorm's dSPM algorithm. The grey shaded areas mark the N1 (light grey) and the P2 (dark grey) time windows. The boxes depict the location of the defined ROIs, with auditory ROIs in blue and visual ROIs in yellow. B) Group effect of the N1 peak mean in the visual cortex: both CI-CHD and CI-SSD users show more activity in the visual cortex compared to NH listeners, regardless of condition. C) Condition effect of the N1 peak mean in the auditory cortex: there is a significantly reduced auditory-cortex activation for AV-V compared to A, indicating multisensory interactions in all groups. D) N1 latency effect in the auditory cortex: Both CI and CI-SSD users show a prolonged N1 latency compared to NH listeners in the auditory cortex, regardless of the condition. This suggests a delayed auditory-cortex activation in CI users, independent of the hearing threshold in the contralateral ear. E) P2 condition effect in the visual cortex: there is a significantly reduced visual-cortex activation for AV-V compared to A, pointing towards multisensory interactions in all groups. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subjects factors for the P2 time window. The three-way mixed ANOVA showed a significant group x map ($F_{4,60} = 3.47$; $p \leq .05$, $\eta p^2 = 0.12$) and a condition x map ($F_{2,60} = 3.41$; $p \leq .05$, $\eta p^2 = 0.03$) interaction. For the group × condition interaction, follow-up t-tests revealed for the NH listeners a significantly enhanced presence of Map C (= P2 topography) compared to CI-SSD users ($t(21) = -3.18$; $p \leq .01$). Vice versa, CI-SSD users showed a significantly enhanced presence of Map E (= P2-like topography) compared to NH listeners ($t(21) = 3.65$; $p \leq .01$), regardless of the condition. Finally, for Map D, there was a significant difference between CI-CHD users and NH individuals ($t(21) = 3.46$; $p \leq .01$), with CI-CHD users showing a more dominant presence of this map compared to NH controls. These results are shown in Fig. 3B. Following the condition × map interaction, follow-up t-tests revealed significant differences between A and AV-V only for Map E ($t(32) = -2.5$; $p \leq .01$). This result suggests that a P2-like topography (Map E) is generated more often for modulated responses (AV-V) compared to unmodulated responses (A),

which is shown in Fig. 3B.

In sum, our results about the *first onset of maps* and the *map presence* at *P2 latency* suggest group-specific topographic differences at P2 latency, with a stronger presence of a conventional P2 topography (Map C) in NH listeners compared to CI-SSD users and a stronger presence of the P2-like topography (Map E) in CI-SSD users compared to NH listeners. Together with the observation that Map D is more present in CI-CHD users than in NH listeners, these results confirm our GMD results, showing a significant group difference between NH listeners and the two CI-user groups for both conditions (A, AV-V) at P2 latency (Fig. 2). Finally, all groups show a P2-like topography (Map E) that is more frequent in the modulated than in the auditory-only condition, which points to alterations in the cortical processing at P2 latency due to the additional visual information in the speech signal.

3.5. Results from ERP source analysis

We conducted a source analysis to further analyse the differences between the three groups, focusing on the auditory and visual cortex activity in both hemispheres. Single-subject source activities for each ROI, condition and group were exported from Brainstorm and were statistically analysed. The source waveforms for the N1 and the P2 are illustrated in Fig. 4A, showing the response in the auditory cortex (N1 peak latency mean: CI-CHD = 141 ms ± 27 ms; CI-SSD = 143 ms ± 28 ms; NH = 122 ms ± 22 ms) and in the visual cortex (N1 peak latency mean: CI-CHD = 143 ms ± 29 ms; CI-SSD = 136 ms ± 28 ms; NH = 136 ms ± 35 ms) for all groups. The peak mean amplitudes and latencies were the dependent variables for the following ANOVA. We performed a mixed-model ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and hemisphere (left, right) as the within-subject factors for each time window of interest (N1, P2) and each ROI (auditory, visual) separately.

Concerning the N1 *peak mean* in the *visual cortex*, the mixed-model ANOVA showed a significant main effect of group ($F_{2,30} = 4.39$; $p \leq .05, \eta p^2 = 0.15$). Post-hoc t-tests confirmed a significant difference between the NH listeners and both CI groups (NH vs. CI-CHD: $t(21) = 3.01$; $p \leq .05$; NH vs. CI-SSD: $t(21) = 3.18$; $p \leq .05$), but no difference between the two CI groups (CI-CHD vs. CI-SSD: $t(21) = -0.24$; $p = 0.81$). Thus, both CI user groups showed more recruitment of the visual cortex compared to NH listeners, regardless of hemisphere and condition (see Fig. 4B).

For the N1 *peak mean* in the *auditory cortex*, the mixed-model ANOVA revealed a significant main effect of hemisphere ($F_{1,29} = 9.46$; $p \leq .005$, $\eta p^2 = 0.11$) and a significant main effect of condition ($F_{1,29} = 34.25$; $p \leq .001, \eta p^2 = 0.11$). Resolving the main effect of hemisphere, follow-up t-tests showed a greater amplitude for the right hemisphere compared to the left hemisphere ($t(63) = -3.66$; $p \leq .001$), regardless of group and condition. Following the main effect of condition, the subsequent t-tests revealed reduced amplitudes for AV-V compared to A ($t(64) = 5.95$; $p \leq .001$), regardless of hemisphere and group, which points to multisensory interaction processes (see Fig. 4C).

For the N1 *peak latency* in the *auditory cortex*, the mixed-model ANOVA identified a significant main effect of group ($F_{2,29} = 4.31$; $p \leq .05, \eta p^2 = 0.13$) and a significant main effect of hemisphere ($F_{1,29} = 5.68$; $p \leq .05, \eta p^2 = 0.02$). Following the main effect of hemisphere, the post-hoc *t*-test revealed a significant difference between the left and the right auditory cortex ($t(63) = 2.20$; $p \leq .05$) with the right hemisphere showing faster latencies compared to the left hemisphere. Resolving the main effect of group, follow-up t-tests revealed a significantly shorter latency of the auditory-cortex response in the NH listeners compared to both CI groups (NH vs. CI-CHD: $t(21) = 5.64$; $p \leq .001$; NH vs. CI-SSD: $t(21) = 6.01$; $p \leq .001$), but no difference between the two CI groups (CI-CHD vs. CI-SSD: $t(21) = -0.20$; $p = 0.80$). Hence, both CI user groups showed a delayed auditory-cortex response compared to NH listeners, regardless of hemisphere and condition (see Fig. 4C). For the N1 *peak latency* in the *visual cortex*, the mixed-model ANOVA did not show any significant main effects or interactions.

Concerning the P2 *peak mean* in the *auditory cortex*, the mixed-model ANOVA found a significant main effect of condition ($F_{1,29} = 11.25$; $p \leq .01, \eta p^2 = 0.04$). Resolving this main effect, the post-hoc t-tests revealed a significant difference between A and AV-V ($t(64) = 2.9$; $p \leq .005$), with A showing greater amplitudes than AV-V, regardless of group and hemisphere. This points to multisensory interaction processes in the auditory cortex at P2 latency.

For the P2 *peak mean* in the *visual cortex*, the mixed-model ANOVA found a significant main effect of condition ($F_{1,30} = 17.06$; $p \leq .001$, $\eta p^2 = 0.06$) as well. Follow-up t-tests revealed a significant difference between A and AV-V ($t(65) = 3.92$; $p \leq .001$), with A showing greater amplitudes than AV-V, regardless of group and hemisphere. This points to multisensory interaction processes in the visual cortex at P2 latency as well (see Fig. 4E).

Regarding the P2 *peak latency* in the *auditory and visual cortices*, the mixed-model ANOVA found neither significant main effects nor significant interactions.

3.6. Correlations

We performed correlations for each CI user group (CI-SSD and CI-CHD), using the Pearson's correlation and the Benjamini-Hochberg (BH) procedure to control for multiple comparisons (Benjamini and Hochberg, 1995). First, we wanted to check whether lip-reading abilities are related to the CI experience and the age at onset of hearing loss (Stropahl et al., 2015; Stropahl and Debener, 2017; Layer et al., 2022). The results revealed a trend for a positive relationship between lip-reading abilities and CI experience (CI-CHD: $r = 0.51$; $p = 0.038$; p corrected = 0.076; CI-SSD: $r = 0.47$; $p = 0.12$; p corrected = 0.147) and a negative relationship between lip-reading abilities and the age of onset of hearing loss for both CI user groups (CI-CHD: $r = -0.84$; $p = 0.001$; p *corrected* = 0.005; CI-SSD: $r = -0.69$; $p = 0.01$; p *corrected* = 0.027). Thus, for both CI user groups it holds that the earlier the onset of hearing impairments, the more pronounced are the lip-reading abilities. Moreover, we aimed to reproduce the relationship between CI experience and the activation in the visual cortex (Giraud et al., 2001c; Layer et al., 2022). The results did not reach a significance level ($CI-CHD: r = 0.51$; $p = 0.1$; p *corrected* = 0.16; $CI-SSD: r = 0.34$; $p = 0.29$; p *corrected* = 0.29).

4. Discussion

In this follow-up study, we used behavioural and EEG measures to investigate audiovisual interactions in CI users with unilateral (CI-SSD) and bilateral (CI-CHD) hearing loss and in a group of NH controls. This study was conducted to extend the results from our previous study comparing CI-CHD users with NH listeners (Layer et al., 2022), by including a third group of participants; namely the CI-SSD users. A subset of our previously reported data was reused and compared to the additional group of CI-SSD users. The inclusion of the additional group of patients significantly extends our previous study because it not only examines the transferability of our previous findings to different CI patient groups, but also provides valuable insights into the influence of individual factors - specifically the hearing ability of the second ear - on audiovisual speech processing in CI users.

At the behavioural level, we confirmed multisensory interactions for all three groups, as evidenced by the shortened response times for the audiovisual condition compared to each of the two unisensory conditions (Fig. 1B) and by the violation of the race model (Fig. 1E). This was in line with the ERP analyses, confirming a multisensory effect for all groups by exhibiting a reduced activation in the auditory and visual cortex for the modulated (AV-V) response compared to the auditory-only (A) response at both the N1 and P2 latencies (Fig. 4C and E,

respectively). In addition to this multisensory effect across all groups, we found group-specific differences. First, specifically the group of CI-CHD users, showed a change of N1 voltage topographies when additional visual information accompanied the auditory information (Fig. 3A), which suggests a particularly strong visual impact on auditory speech processing in CI users with bilateral hearing loss. Second, both groups of CI users revealed a delayed auditory-cortex activation (Fig. 4D), enhanced lip-reading abilities (Fig. 1D) and stronger visual-cortex activation (Fig. 4B) when compared to the NH controls. Thus, the current results extend the results of our previous study (Layer et al., 2022) by showing distinct multisensory processes not only between NH listeners and CI users in general, but even between CI users with unilateral (CI-SSD) and bilateral (CI-CHD) hearing loss.

4.1. Behavioural multisensory integration in all groups

The behavioural results revealed that both the NH listeners and the two CI user groups had faster reaction times for audiovisual syllables than for unisensory (auditory-alone, visual-alone) syllables (Fig. 1; Table 2). No difference was found between the auditory and visual conditions. Hence, all groups exhibited a clear redundant signals effect for audiovisual syllables, implying that the benefit of cross-modal input is comparable between the CI user groups and NH listeners on a behavioural level (Laurienti et al., 2004; Schierholz et al., 2015; Layer et al., 2022), at least when considering syllables that are combined with a talking head. The violation of the race model for each group (CI-CHD, CI-SSD, NH) suggests that multisensory integration was the cause for the observed redundant signals effect in both CI user groups and NH listeners. However, the behavioural responses of the CI users were not slower compared to the NH listeners, even though the signal provided by the CI is known to be limited in comparison to a natural hearing experience (Drennan and Rubinstein, 2008). The observation of comparable response times in CI users can be explained by the fact that there were only two syllables, and that the difficulty of the task was correspondingly low. Compatible with this, all groups were equally able to perform the task, and the subjective rating of the listening effort showed no difference between the groups.

One would assume that the CI users might be better and faster at identifying the purely visual syllables due to results from previous studies with congenitally deaf individuals and CI users, showing visual enhancements, in particular visually induced activation in the auditory cortex (Bottari et al., 2014; Finney et al., 2003; Hauthal et al., 2014; Bavelier and Neville, 2002; Heimler et al., 2014; Sandmann et al., 2012). This cross-modal activation seems to be driven by auditory deprivation and might form the neural basis for specific superior visual abilities (Lomber et al., 2010). Importantly, auditory impairment is not only experienced in CI users before receiving a CI, but also after the implantation when only a limited auditory input is provided by the CI. Thus, it is not surprising that CI users reveal compensatory visual strategies, such as enhanced lip-reading abilities, in order to overcome the limited CI signal (Rouger et al., 2007; Schreitmüller et al., 2018; Stropahl et al., 2015; Stropahl and Debener, 2017). Our results extend previous observations of enhanced visual abilities in CI users by showing that not only CI-CHD users, but surprisingly also CI-SSD users demonstrate a better lip-reading ability when compared to NH listeners. Importantly, the lip-reading ability was comparable between the two patient groups, and both groups showed a positive correlation with the age of the onset of hearing loss, indicating that an earlier onset of hearing loss triggers improved behavioural visual abilities. Our results demonstrate that this visual improvement develops across different groups of CI patients, independent of the hearing abilities of the contralateral ear.

However, behavioural visual improvements in CI users seem to be stimulus- and task-specific, as indicated by our finding that the two CI user groups showed comparable behavioural results to NH listeners in the speeded response task. Our finding is consistent with previous studies, using a speeded response task with simple tones and white discs as auditory and visual stimuli, respectively (Schierholz et al., 2015, 2017). It seems that in our study the task with the basic stimuli and the two syllables was too easy, leading to ceiling effects in all groups. This estimation is in line with our observation that the perceived exertion effort was comparable between all three groups. Importantly, behavioural group differences have well been reported in a previous study using more complex stimuli presented in the context of a difficult recognition paradigm, showing an enhanced audiovisual gain in CI users when compared to NH listeners (Radecke et al., 2022). This is consistent with the view that behavioural advantages due to additional visual information in CI users are task- and stimulus-selective, and that they become evident under specific circumstances, for instance in conditions with semantic information (Moody-Antonio et al., 2005; Rouger et al., 2008; Tremblay et al., 2010; Radecke et al., 2022). Thus, future studies should use linguistically complex stimuli, such as words or sentences presented in auditory, visual and audiovisual conditions, in order to better understand the behavioural advantages for visual and audiovisual speech conditions in CI users compared to NH individuals.

4.2. Electrophysiological correlates of multisensory speech perception

Similar to the behavioural data, we also discovered commonalities among groups at the ERP level. Nonetheless, group differences were found as well, which will be discussed in the following sections.

4.2.1. Group similarities in multisensory speech processing

Similar to the behavioural results, we found evidence for multisensory effects in the ERP responses. The topographic clustering analysis with subsequent single-subject fitting confirmed multisensory interactions for both CI user groups and NH individuals by revealing an increase in *P2-like topographies* for the modulated ERPs (AV-V) compared to the purely auditory condition (A). This observation points to a visual modulation of the auditory ERPs in the two CI user groups as well as in the NH individuals.

These findings are supported by the source analysis. Specifically, we investigated the visual impact on the auditory cortical response by comparing the ERPs of the auditory condition (A) with the visually modulated ERPs (AV-V). A difference between these two conditions can be seen as evidence for non-linear multisensory interactions (Besle et al., 2004; Murray et al., 2005; Cappe et al., 2010; Foxe et al., 2000; Molholm et al., 2002). As previous studies have shown, a difference between two conditions can either be superadditive ($A < AV - V$; e.g. Schierholz et al., 2015) or subadditive ($A > AV - V$; e.g. Cappe et al., 2010) (Stevenson et al., 2014). Our source analyses confirmed the multisensory interactions in all groups, both within the N1 and the P2 time windows by demonstrating a subadditive effect, indicated by a reduced activation in the auditory cortex for the modulated (AV-V) compared to the auditory-only (A) condition. A similar subadditive effect was even observed for the visual cortex at P2 latency in all groups. In sum, these findings are highly consistent with the results from other ERP studies (e.g. Cappe et al., 2010) as well as fMRI studies (e.g. Martuzzi et al., 2007), showing multisensory interactions in both the auditory and visual cortex and confirming the behavioural results of multisensory integration for all groups.

4.2.2. Group differences in multisensory speech processing

As measured by global field power (GFP), the signal strength of cortical responses did not differ between the two CI user groups and the NH listeners. Nevertheless, for the response topography, as quantified by the global map dissimilarity (GMD), we detected various differences between CI-CHD, CI-SSD and NH listeners for both the auditory-only condition (A) and the modulated response (AV-V). Thus, group differences were not caused by signal strength but rather by differences in their electric field topographies and by extension the configuration of active brain networks. To ascertain whether distinct topographies were

caused by a simple latency shift of the ERP responses and/or by variable configurations of the neural sources, we conducted various analyses, in particular hierarchical clustering, single-subject fitting and source analysis.

Within the N1 time window, hierarchical clustering and single-subject fitting showed specifically for the CI-CHD users – but not for CI-SSD users and not for NH listeners – a greater presence of N1 topographies (Map B) for modulated responses (AV-V) compared to purely auditory responses (A). The difference in N1 topography between CI-CHD users and NH listeners was already reported in our previous study (Layer et al., 2022), which was now expanded by including a group of CI-SSD users. Interestingly, the NH listeners and CI-SSD users showed no differences in map presence between auditory and modulated responses, indicating that these two groups use similar neural processes to evaluate the audiovisual speech stimuli. An increasing presence of a conventional N1 topography (Map B) specifically for the modulated condition indicates alterations in audiovisual processing and a multisensory benefit for CI-CHD users, when additional visual information is present. This modulation in CI-CHD users reflects most likely a strategy that they develop to compensate for the limited CI input. Interestingly, this modulation was not detected in CI-SSD users, which leads to the assumption that CI-CHD have a higher benefit from additional visual input compared to CI-SSD users. This group-specific effect may be due to the fact that CI-SSD users have a NH ear on the contralateral side which might serve as the main communication channel. Therefore, it is not surprising that CI-SSD users and NH listeners show comparable topographic results, at least within the N1 latency range. Previous research comparing different groups of CI users is limited. Nevertheless, first evidence of differences in speech-in-noise performance between CI-SSD users and bimodal CI users (CI on one ear and hearing aid on the contralateral ear) was reported (Williges et al., 2019). On the other hand, differences between CI-SSD users and bilateral CI users (both ears fitted with CIs) were observed in situations with multiple concurrent speakers (Bernstein et al., 2016). As a result, we expected group differences to emerge not only for auditory stimulation but also for audiovisual stimulation. However, as far as we are aware, this has not been investigated yet. Our results therefore indeed confirm first indications of different processing strategies among different CI user groups.

Given that our results are restricted to the processing of syllables, we propose that further studies should use more complex speech stimuli in order to refine and deepen the current findings about different groups of CI users. Similar to the current analyses, future studies should apply electrical neuroimaging (Michel et al., 2009), including topographic and ERP source analysis, given that it is a powerful approach to investigate multisensory interactions (Stevenson et al., 2014). It is likely that the testing of CI-CHD in more complex speech conditions, in particular with semantic information, results in even more enhanced cortical audiovisual interactions (Radecke et al., 2022), as can be assumed based on previous behavioural results (Moody-Antonio et al., 2005; Rouger et al., 2008; Tremblay et al., 2010). We speculate that in these demanding stimulus conditions, an increase in audiovisual interactions may become even detectable in the group CI-SSD users.

Apart from this modulation effect, which was specific for CI-CHD users, we found generally delayed cortical responses for both CI user groups compared to NH listeners, which was consistently reflected in various analyses. On the sensor level, ERP data revealed that both CI-CHD users and CI-SSD users had a prolonged N1 latency compared to NH listeners for both the modulated and the auditory-only responses. By comparing the first onset of Map B (N1 topography) between the three groups of participants, hierarchical clustering and single-subject fitting analyses confirmed the delayed onset of the N1 topography in the two groups of CI users. Similarly, the source analysis revealed a delayed N1 response in the auditory cortex for both CI-CHD and CI-SSD users compared to NH listeners. In fact, since the signal does not have to pass through the middle and inner ear due to direct stimulation of auditory nerve fibres, one might assume that the time until the electrical signal reaches the auditory nerve with a CI is faster than in NH listeners. Interestingly, Schierholz et al. (2017) discovered delayed cortical responses not only in patients with a CI, but also in patients with an auditory brainstem implant (ABI; direct stimulation of the cochlear nucleus) and an auditory midbrain implant (AMI; direct stimulation of the inferior colliculus). Despite the fact that central auditory implants bypass more structures than CIs, ABI and AMI patients showed even more delayed cortical and poorer behavioural responses when compared to CI patients, appearing to be specifically related to the insufficient input provided by central auditory implants (Schierholz et al., 2017). Analogously, delayed N1 ERP responses in CI users are likely to reflect difficulties in processing speech sounds with the CI compared to natural hearing. In line with this, previous research with NH listeners has shown that difficult acoustic listening conditions, such as speech in background noise, cause a delay in the N1 response (Billings et al., 2011; Finke et al., 2016a). Furthermore, our findings of slowed cortical N1 ERPs in CI users are consistent with the findings of several other studies which used auditory stimuli of varying acoustic complexity, and which showed delayed N1 latency and poorer auditory discrimination ability in CI-CHD users compared to NH listeners (Beynon et al., 2005; Finke et al., 2016a; Sandmann et al., 2009, 2015; Senkowski et al., 2014). This is also an observation that has been reported in studies with CI-SSD users comparing the cortical responses from the CI ear and the NH ear (Finke et al., 2016b; Weglage et al., 2022).

To sum up, the results on the N1 ERPs revealed several group differences. First, the hierarchical clustering and single-subject fitting analyses showed a *condition effect* specifically for the group of CI-CHD users, showing distinct patterns of N1 topography between the modulated auditory response (AV-V) and the auditory-only response (A; Fig. 3A). However, for both the CI-SSD users and the NH listeners, such a modulation was not given, which leads to the conclusion that these two groups do not benefit from additional visual input to the same extent as CI-CHD users. CI-CHD users seem to have a particularly strong visual impact on auditory speech processing, allowing these individuals to compensate the limited CI input. By contrast, CI-SSD users appear to be less dependent on additional visual input, probably due to the fact that they can rely on the contralateral ear, which is normal-hearing.

Apart from the CI-CHD-specific condition effect, however, our results revealed a *map dissimilarity at N1 latency* between the three groups, at least for the auditory condition (Fig. 2A). The results from the topographic and source analyses suggest that these group differences can at least partially be explained by a delayed N1 auditory-cortex response in both CI-CHD and CI-SSD users when compared to NH listeners. However, as discussed in the following section (4.2.3), the results from the source analyses indicate that this map dissimilarity at N1 latency is also caused by a different configuration of neural sources. Both CI user groups showed an additional activation in the visual cortex compared to NH listeners (see section 4.2.3 for more details).

Within the P2 time window, the hierarchical clustering and single-subject fitting analyses showed a condition effect (independent of group; Fig. 3B) and some group effects (independent of condition; Fig. 3B). The *condition* effect, which revealed an enhanced map presence of the P2-like topography (Map E) for the modulated (AV-V) compared to the auditory response (A), indicates a visual modulation of the auditory ERPs in all groups of participants. Regarding the *group* effects, we found a greater map presence of the P2 topography (Map C) for NH individuals compared to CI-SSD users, and a greater map presence of a P2-like topography (Map E) for CI-SSD users compared to NH listeners. This difference is in line with the results from the GMD observations, pointing to a difference in response topography at the P2 time window (Fig. 2). At the same latency window, we also found a difference between CI-CHD users and NH listeners for the Map D, showing a greater presence of this topography in CI-CHD users than NH listeners. This difference was also confirmed by the GMD analysis, showing a difference in response topography at the P2 time window (Fig. 2). Importantly, no significant differences between CI-CHD users and CI-SSD users

were detected in the topographic analysis at P2 latency, which corresponds to the lack of differences in the GMD analysis. These results however are not easy to interpret. One may speculate that the two CI user groups do not differ statistically at P2 latency but exhibit distinct topographic patterns when compared to NH listeners. Therefore, a source analysis is unlikely to yield clear results revealing group-specific cortical activation patterns. Indeed, the source analysis revealed no group differences in auditory or visual cortex activation. Nevertheless, the source analysis provided evidence for audiovisual interactions by identifying reduced amplitudes for the modulated condition (AV-V) compared to the purely auditory condition (A), independent of the groups (see section 4.1). Taken together, we conclude that CI users and NH listeners recruit the auditory and visual cortices similarly during auditory and audiovisual speech processing, at least at the P2 latency range. Future research is required to determine if group-specific differences within the P2 time window emerge with other stimuli/tasks and whether they appear beyond the visual and auditory sensory systems, as in frontal or other temporal areas (Campbell and Sharma, 2013; Giraud et al., 2001a).

4.2.3. Cross-modal plasticity in the visual cortex

The results from the source analysis within the N1 time window revealed a greater recruitment of the visual cortex for both the CI-CHD users and CI-SSD users when compared to NH listeners. This cross-modal activation of the visual cortex is in line with previous research findings from CI-CHD users, reporting that cortical alterations are not limited to the auditory cortex (Campbell and Sharma, 2013, 2016; Chen et al., 2016; Giraud et al., 2001a,b) but seem to extend to the visual and even inferior frontal areas (Rouger et al., 2012). As far as we are aware, this is the first study to show that auditory-induced activation in the visual cortex is not restricted to CI-CHD users, as shown in our previous study (Layer et al., 2022) but is also present in CI-SSD users.

It has been suggested that the additional recruitment of the visual cortex in CI users is a way to compensate the limited auditory input delivered through the CI (Doucet et al., 2006; Giraud et al., 2001c; Strelnikov et al., 2010, 2013). In the PET (positron emission tomography) study of Giraud and colleagues (Giraud et al., 2001c), a greater auditory activation in the visual cortex was reported for unilaterally-implanted CI users (with bilateral hearing loss) in comparison to NH controls for meaningful sounds. The authors found an association between this visual recruitment with longer CI experience and pronounced lip-reading abilities (Giraud et al., 2001c). By now, cross-modal recruitment of the visual cortex for processing auditory stimuli has been observed in CI users for syllables, words, environmental sounds, pure tones, and reversed words (Chen et al., 2016; Giraud et al., 2001b; Giraud et al., 2001c). This visual cortex activation when processing auditory stimuli may represent an enhanced synergy of the auditory and visual modalities in CI users. This might be beneficial for speech comprehension recovery following cochlear implantation, when CI users must learn to match novel auditory speech inputs with corresponding visual speech cues (Strelnikov et al., 2013). Consistent with this view, our previous study showed a positive correlation between CI experience and visual-cortex activation, indicating that CI users increasingly recruit the visual cortex with increasing duration of CI usage (Layer et al., 2022).

The present study extends previous findings by showing visual recruitment for auditory (and audiovisual) syllables not only in CI-CHD users but also in CI-SSD users. This observation in CI-SSD users may appear surprising because these individuals have at least one ear on the contralateral side on which to rely. However, in natural conditions with bilateral stimulation, the CI-SSD users' auditory performance is still impaired compared to normal hearing patients (Dorman et al., 2015; Ludwig et al., 2021), particularly in difficult listening conditions, which may explain why these individuals develop compensatory strategies, as indicated by enhanced lip-reading skills and cross-modal recruitment of the visual cortex. Taken together, our results suggest that auditory-induced activation of the visual cortex is independent of the hearing abilities on the contralateral ear when being stimulated over the CI ear. However, future studies are required to systematically compare the stimulation of the CI ear alone, the contralateral ear alone and both ears together. The measured cortical activation differences would provide a clearer picture of how much of the large variability in speech recognition ability in different CI users can be attributed to the contralateral ear. Further, we suggest that future studies should further split the CI groups into bimodal (CI – HA), bilateral (CI–CI) and SSD (CI–NH) CI users and use more ecologically valid stimuli such as complex speech stimuli, and different task conditions (e.g. passive vs. active) in order to deliver a portrait of the characteristics for each group in terms of audiovisual speech processing and cortical reorganisation. This could allow for developing methods for accelerating and improving auditory rehabilitation after implantation that are specific for each CI user group.

5. Limitations

An important limitation of this study is the small sample size of n = 11 per group. A small sample size is common in the scarce literature on CI-SSD users, because this patient group was not implanted a few years back due to the intact NH ear. Only after proving the benefits of an implantation (improved hearing abilities compared to other techniques such as bone-anchored hearing aids (BAHA) or contralateral routing of signal (CROS) (Arndt et al., 2011, 2017); better speech intelligibility, sound localisation and quality of life (Kitterick et al., 2015)), CI-SSD users were increasingly fitted with a CI. Consequently, results with small sample sizes always must be interpreted with caution. Future studies, however, should aim to increase the number of participants in order to improve statistical power and apply more extensive statistical correction methods, as suggested by Cramer et al. (2016).

One might ask the question whether the presented results are transferrable to other audiovisual stimuli or whether these are restricted to language-specific stimuli. We assume that the significant audiovisual benefit observed in our CI users was strongly driven by the linguistic property of the presented stimuli, and we hypothesise that this audiovisual benefit may even be more pronounced for more difficult linguistic stimuli (words or sentences). Nevertheless, when solely focusing on the additional recruitment of the visual cortex, as observed in the present and our recent results (Layer et al., 2022), previous studies have suggested that this effect is not restricted to purely linguistic tasks. This effect has been shown for syllables, words and environmental sounds (Giraud et al., 2001c). Another study (Chen et al., 2016) extended these results to pure tones and reversed words, showing that intelligibility is not necessary for eliciting auditory-induced recruitment of the visual cortex. However, these previous studies do not provide answers concerning processing differences and similarities between CI user groups, which is the novelty of this study. We hypothesise that further and more pronounced differences between CI-CHD users and CI-SSD users will become evident for more difficult linguistic stimuli (words/sentences vs. syllables) and more difficult task conditions (semantic processing vs. discrimination of syllables). Importantly, future studies should use the same experimental settings (i.e. the same paradigm) for different types of stimuli. They should compare cortical response patterns between non-linguistic stimuli (e.g. basic and environmental sounds) and linguistic stimuli (e.g., syllables and words), to see if our findings are transferable to both non-linguistic and more difficult linguistic stimulus conditions within the same patients.

6. Conclusion

The present study applied electrical neuroimaging, including topographic and source analysis, to investigate whether the cortical processing of audiovisual syllables is different between CI-CHD users and CI-SSD users. These two CI user groups were also compared to NH listeners. This study expands on our previous study that compared CI-CHD

users to NH listeners by reusing these data to include an additional group of CI users. Our findings showed a clear multisensory effect in both CI user groups and NH listeners, as evidenced by faster behavioural response times and reduced auditory-cortex activation at N1 and P2 latencies. Furthermore, we discovered a condition effect for the N1 ERP topography in the group of CI-CHD users only, demonstrating a distinct pattern of N1 topography between the auditory-only (A) and the modulated auditory response (AV-V). This finding indicates a strong visual influence on auditory speech processing especially in CI-CHD users, allowing these patients to compensate for the limited CI input. Finally, we discovered that both CI user groups showed a delay in the auditory-cortex response at N1 latency, indicating difficulties in cortical processing of the CI's limited signal. Nonetheless, we found that both CI user groups showed pronounced lip-reading abilities and an additional recruitment of the visual cortex compared to NH listeners. This finding extends previous results by demonstrating that this cross-modal visual activation is also present in CI-SSD users, who have an intact NH ear on the contralateral side. One may speculate that these cortical alterations allow the CI users to combine multisensory information, to refine expectations and to sharpen perception. We conclude that an auditory-induced activation in the visual cortex is independent of the hearing threshold and supply on the contralateral ear when being stimulated via the CI ear. This auditory-induced activation in the visual cortex is an important and insightful similarity between the two CI user groups. Overall, these results confirm existing differences in multisensory processing both when comparing NH listeners and CI users and when comparing different groups of CI users. This emphasises the importance of developing individual rehabilitation methods tailored to different groups of CI users.

Funding

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft; DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and by the Köln Fortune Program/Faculty of Medicine, University of Cologne [278/2020] to PS. MMM is supported by the Swiss National Science Foundation (169206).

CRediT authorship contribution statement

Natalie Layer: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration. **Anna Weglage:** Software, Validation, Formal analysis, Investigation, Writing – review & editing. **Verena Müller:** Conceptualization, Writing – review & editing. **Hartmut Meister:** Conceptualization, Software, Resources, Writing – review & editing. **Ruth Lang-Roth:** Conceptualization, Writing – review & editing. **Martin Walger:** Conceptualization, Writing – review & editing. **Micah M. Murray:** Software, Formal analysis, Validation, Writing – review & editing. **Pascale Sandmann:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition.

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.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and Köln Fortune [278/2020] for providing funding for this research. Also, we thank Pauline Burkhardt and Lina Wiesel for their help in data collection and Moritz Layer for his support with the implementation of the analysis. Finally, we thank all participants for taking part in this study.

Appendix A. Peer Review Overview and Supplementary data

A Peer Review Overview and (sometimes) Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.crneur.2022.100059.

References

Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., 2006. Task-modulated "what" and "where" pathways in human auditory cortex. Proc. Natl. Acad. Sci. USA 103 (39), 14608–14613.
Allen, J.B., Hall, J., Jeng, P., 1990. Loudness growth in 1/2-octave bands (LGOB)—a procedure for the assessment of loudness. J. Acoust. Soc. Am. 88 (2), 745–753.
Arndt, S., Aschendorff, A., Laszig, R., Beck, R., Schild, C., Kroeger, S., Ihorst, G., Wesarg, T., 2011. Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. Otol. Neurotol. 32 (1), 39–47.
Arndt, S., Laszig, R., Aschendorff, A., Hassepass, F., Beck, R., Wesarg, T., 2017. Cochleaimplantatversorgung bei einseitiger Taubheit oder asymmetrischem Hörverlust. HNO 65 (7), 586–598.
Barth, D.S., Goldberg, N., Brett, B., Di, S., 1995. The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. Brain Res. 678 (1–2), 177–190.
Bavelier, D., Neville, H.J., 2002. Cross-modal plasticity: where and how? Nat. Rev. Neurosci. 3 (6), 443–452.
Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7 (6), 1129–1159.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B 57 (1), 289–300.
Bernstein, J.G., Goupell, M.J., Schuchman, G.I., Rivera, A.L., Brungart, D.S., 2016. Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. Ear Hear. 37 (3), 289.
Besle, J., Fort, A., Delpuech, C., Giard, M.-H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. Eur. J. Neurosci. 20 (8), 2225–2234.
Beynon, A., Snik, A., Stegeman, D., Van den Broek, P., 2005. Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients, 01 J. Am. Acad. Audiol. 16, 042, 053.
Biasiucci, A., Franceschiello, B., Murray, M.M., 2019. Electroencephalography. *Current Biology* 29 (3), R80–R85.
Billings, C.J., Bennett, K.O., Molis, M.R., Leek, M.R., 2011. Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. Ear Hear. 32 (1), 53.
Bönitz, H., Kopp, B., Büchner, A., Lunner, T., Lyxell, B., Finke, M., 2018. Event-related neuronal responses to acoustic novelty in single-sided deaf cochlear implant users: initial findings. Clin. Neurophysiol. 129 (1), 133–142.
Bosnyak, D.J., Eaton, R.A., Roberts, L.E., 2004. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 hz amplitude modulated tones. Cerebr. Cortex 14 (10), 1088–1099.
Bottari, D., Heimler, B., Caclin, A., Dalmolin, A., Giard, M.-H., Pavani, F., 2014. Visual change detection recruits auditory cortices in early deafness. Neuroimage 94, 172–184.
Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., Röder, B., 2020. EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. Neuroimage 223, 117315.
Brünecke, I., Hölsken, S., Kessler, J., 2018. Der DemTect Eye+ Ear–Neues kognitives Screeningverfahren bei schwerhörigen Menschen mit Demenzverdacht. Z. für Audiol. 57 (3), 121.
Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput. Intell. Neurosci. 2011.
Buechner, A., Brendel, M., Lesinski-Schiedat, A., Wenzel, G., Frohne-Buechner, C., Jaeger, B., Lenarz, T., 2010. Cochlear implantation in unilateral deaf subjects associated with ipsilateral tinnitus. Otol. Neurotol. 31 (9), 1381–1385.
Campbell, J., Sharma, A., 2013. Compensatory changes in cortical resource allocation in adults with hearing loss. Front. Syst. Neurosci. 7, 71.
Campbell, J., Sharma, A., 2016. Visual cross-modal re-organization in children with cochlear implants. PLoS One 11 (1), e0147793.
Cappe, C., Thut, G., Romei, V., Murray, M.M., 2010. Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. J. Neurosci. 30 (38), 12572–12580.
Chen, L.-C., Sandmann, P., Thorne, J.D., Bleichner, M.G., Debener, S., 2016. Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fNIRS. Neural Plast. 4382656–4382656.
Cramer, A.O., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R.P., et al., 2016. Hidden multiplicity in exploratory multiway ANOVA: prevalence and remedies. Psychonomic Bull. Rev. 23 (2), 640–647.
Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115 (4), 732–744.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26 (1), 55–67.
Debener, S., Hine, J., Bleeck, S., Eyles, J., 2008. Source localization of auditory evoked potentials after cochlear implantation. Psychophysiology 45 (1), 20–24.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21.
Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53 (1), 1–15.
Dong, L., Foo, S.W., Lian, Y., 2003. Modeling continuous visual speech using boosted viseme models. In: Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint, 3, pp. 1394–1398.
Dorman, M.F., Zeitler, D., Cook, S.J., Loiselle, L., Yost, W.A., Wanna, G.B., Gifford, R.H., 2015. Interaural level difference cues determine sound source localization by single-sided deaf patients fit with a cochlear implant. Audiol. Neurotol. 20 (3), 183–188.
Doucet, M., Bergeron, F., Lassonde, M., Ferron, P., Lepore, F., 2006. Cross-modal reorganization and speech perception in cochlear implant users. Brain 129 (12), 3376–3383.
Drennan, W.R., Rubinstein, J.T., 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. J. Rehabil. Res. Dev. 45 (5), 779.
Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H.E., Knight, R.T., Deouell, L.Y., 2019. Direct evidence for prediction signals in frontal cortex independent of prediction error. Cerebr. Cortex 29 (11), 4530–4538.
Fagel, S., Clemens, C., 2004. An articulation model for audiovisual speech synthesis—determination, adjustment, evaluation. Speech Commun. 44 (1–4), 141–154.
Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., Sandmann, P., 2016a. On the relationship between auditory cognition and speech intelligibility in cochlear implant users: an ERP study. Neuropsychologia 87, 169–181.
Finke, M., Sandmann, P., Bönitz, H., Kral, A., Büchner, A., 2016b. Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. Audiol. Neurotol. 21 (5), 305–315.
Finney, E.M., Clementz, B.A., Hickok, G., Dobkins, K.R., 2003. Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. Neuroreport 14 (11), 1425–1427.
Foxe, J.J., Morocz, I.A., Murray, M.M., Higgins, B.A., Javitt, D.C., Schroeder, C.E., 2000. Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. Cognit. Brain Res. 10 (1–2), 77–83.
Giraud, A.L., Price, C.J., Graham, J.M., Frackowiak, R.S., 2001a. Functional plasticity of language-related brain areas after cochlear implantation. Brain 124 (7), 1307–1316.
Giraud, A.L., Truy, E., Frackowiak, R., 2001b. Imaging plasticity in cochlear implant patients. Audiol. Neurotol. 6 (6), 381–393.
Giraud, A.-L., Price, C.J., Graham, J.M., Truy, E., Frackowiak, R.S., 2001c. Cross-modal plasticity underpins language recovery after cochlear implantation. Neuron 30 (3), 657–664.
Glennon, E., Svirsky, M.A., Froemke, R.C., 2020. Auditory cortical plasticity in cochlear implant users. Curr. Opin. Neurobiol. 60, 108–114.
Godey, B., Schwartz, D., De Graaf, J.B., Chauvel, P., Liegeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin. Neurophysiol. 112 (10), 1850–1859.
Gordon, K.A., Jiwani, S., Papsin, B., 2013. Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf. Front. Psychol. 4, 719.
Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. Biomed. Eng. Online 9 (1), 1–20.
Green, K.M., Julyan, P.J., Hastings, D.L., Ramsden, R.T., 2005. Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. Hear. Res. 205 (1–2), 184–192.
Hahlbrock, K.-H., 1970. Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet. Georg Thieme Verlag.
Hansen, P., Kringelbach, M., Salmelin, R., 2010. MEG: an Introduction to Methods. Oxford university press.
Hari, R., Pelizzone, M., Mäkelä, J.P., Hällström, J., Leinonen, L., 1987. Neuromagnetic responses of the human auditory cortex to on-and offsets of noise bursts. Audiology 26, 31–43.
Hauthal, N., Thorne, J.D., Debener, S., Sandmann, P., 2014. Source localisation of visual evoked potentials in congenitally deaf individuals. Brain Topogr. 27 (3), 412–424.
Heimler, B., Weisz, N., Collignon, O., 2014. Revisiting the adaptive and maladaptive effects of crossmodal plasticity. Neuroscience 283, 44–63.
Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C., 2014. Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. Audiol. Neurotol. 19 (Suppl. 1), 21–26.
Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37 (2), 163–178.
Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. USA 97 (22), 11793–11799.
Kitterick, P.T., Lucas, L., Smith, S.N., 2015. Improving health-related quality of life in single-sided deafness: a systematic review and meta-analysis. Audiol. Neurotol. 20 (Suppl. 1), 79–86.
Kral, A., Hubka, P., Heid, S., Tillein, J., 2013. Single-sided deafness leads to unilateral aural preference within an early sensitive period. Brain 136 (1), 180–193.
Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H., Wallace, M.T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. Exp. Brain Res. 158 (4), 405–414.
Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M., Sandmann, P., 2022. The Timecourse of Multisensory Speech Processing in Unilaterally Stimulated Cochlear Implant Users Revealed by ERPs. Neuroimage: Clinical, 102982.
Lazard, D.S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P.H., Skarzynski, H., Schauwers, K., O'Leary, S., Mawman, D., Maat, B., Kleine-Punte, A., Huber, A.M., Green, K., Govaerts, P.J., Fraysse, B., Dowell, R., Dillier, N., Burke, E., Beynon, A., Bergeron, F., Başkent, D., Artières, F., Blamey, P.J., 2012b. Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. PLoS One 7, e48739.
Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48 (6), 609–621.
Lin, F.-H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hämäläinen, M.S., 2006. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. Neuroimage 31 (1), 160–171.
Lomber, S.G., Meredith, M.A., Kral, A., 2010. Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. Nat. Neurosci. 13 (11), 1421–1427.
Ludwig, A.A., Meuret, S., Battmer, R.-D., Schönwiesner, M., Fuchs, M., Ernst, A., 2021. Sound localization in single-sided deaf participants provided with a cochlear implant. Front. Psychol. 12.
Martuzzi, R., Murray, M.M., Michel, C.M., Thiran, J.P., Maeder, P.P., Clarke, S., Meuli, R. A., 2007. Multisensory interactions within human primary cortices revealed by BOLD dynamics. Cerebr. Cortex 17 (7), 1672–1679.
Massaro, D.W., Light, J., 2004. Using Visible Speech to Train Perception and Production of Speech for Individuals with Hearing Loss.
Meister, H., Fuersen, K., Schreitmueller, S., Walger, M., 2016. Effect of acoustic fine structure cues on the recognition of auditory-only and audiovisual speech. J. Acoust. Soc. Am. 139 (6), 3116–3120.
Merzenich, M.M., Van Vleet, T.M., Nahum, M., 2014. Brain plasticity-based therapeutics. Front. Hum. Neurosci. 8, 385.
Micco, A.G., Kraus, N., Koch, D.B., McGee, T.J., Carrell, T.D., Sharma, A., Nicol, T., Wiet, R.J., 1995. Speech-evoked cognitive P300 potentials in cochlear implant recipients. Am. J. Otol. 16 (4), 514–520.
Michel, C.M., Koenig, T., 2018. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. Neuroimage 180, 577–593.
Michel, C.M., Murray, M.M., 2012. Towards the utilization of EEG as a brain imaging tool. Neuroimage 61 (2), 371–385.
Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., de Peralta, R.G., 2004. EEG source imaging. Clin. Neurophysiol. 115 (10), 2195–2222.
Michel, C.M., Koenig, T., Brandeis, D., Gianotti, L.R., Wackermann, J., 2009. Electrical Neuroimaging. Cambridge University Press.
Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. Cognit. Psychol. 14 (2), 247–279.
Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognit. Brain Res. 14 (1), 115–128.
Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer Jr., E.T., Fisher, L., Bernstein, L.E., 2005. Improved speech perception in adult congenitally deafened cochlear implant recipients. Otol. Neurotol. 26 (4), 649–654.
Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., et al., 2005. Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cerebr. Cortex 15 (7), 963–974.
Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20 (4), 249–264.
Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24 (4), 375–425.
Obleser, J., Elbert, T., Lahiri, A., Eulitz, C., 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. Cognit. Brain Res. 15 (3), 207–213.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the edinburgh inventory. Neuropsychologia 9 (1), 97–113.
Park, L.R., Preston, E., Noxon, A.S., Dillon, M.T., 2021. Comparison of test methods to assess the implanted ear alone for pediatric cochlear implant recipients with single-sided deafness. Cochlear Implants Int. 1–8.
Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72 (2), 184–187.
Prince, P., Paul, B.T., Chen, J., Le, T., Lin, V., Dimitrijevic, A., 2021. Neural correlates of visual stimulus encoding and verbal working memory differ between cochlear implant users and normal-hearing controls. Eur. J. Neurosci. 54(3), 5016–5037.
R Core Team, 2020. R: a language and environment for statistical computing. R Found. Stat. Comput. https://www.R-project.org/.
Raab, D.H., 1962. Division of psychology: statistical facilitation of simple reaction times. Trans. N. Y. Acad. Sci. 24 (5 Series II), 574–590.
Radecke, J.O., Schierholz, I., Kral, A., Lenarz, T., Murray, M.M., Pascale, S., 2022. Distinct Multisensory Perceptual Processes Guide Enhanced Auditory Recognition Memory in Older Cochlear Implant Users. NeuroImage: Clinical, 102942.

Ratcliff, R., 1979. Group reaction time distributions and an analysis of distribution statistics. Psychol. Bull. 86 (3), 446.
Ross, B., Tremblay, K., 2009. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. Hear. Res. 248 (1–2), 48–59.
Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P., 2007. Evidence that cochlear-implanted deaf patients are better multisensory integrators. Proc. Natl. Acad. Sci. USA 104 (17), 7295–7300.
Rouger, J., Fraysse, B., Deguine, O., Barone, P., 2008. McGurk effects in cochlear-implanted deaf subjects. Brain Res. 1188, 87–99.
Rouger, J., Lagleyre, S., Démonet, J.-F., Fraysse, B., Deguine, O., Barone, P., 2012. Evolution of crossmodal reorganization of the voice area in cochlear-implanted deaf patients. Hum. Brain Mapp. 33 (8), 1929–1940.
Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., Hugdahl, K., Meyer, M., 2009. Evaluation of evoked potentials to dyadic tones after cochlear implantation. Brain 132 (7), 1967–1979.
Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Debener, S., Jancke, L., Meyer, M., 2010. Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. Clin. Neurophysiol. 121 (12), 2070–2082.
Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R.D., Marcar, V.L., Jäncke, L., Debener, S., 2012. Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. Brain 135 (2), 555–568.
Sandmann, P., Plotz, K., Hauthal, N., Vos, M. de, Schönfeld, R., Debener, S., 2015. Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. Clin. Neurophysiol. 126 (3), 594–607.
Schiel, F., 1999. Automatic phonetic transcription of non-prompted speech. Open Access LMU.
Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., Büchner, A., Dengler, R., Sandmann, P., 2015. Enhanced audio–visual interactions in the auditory cortex of elderly cochlear-implant users. Hear. Res. 328, 133–147.
Schierholz, I., Finke, M., Kral, A., Büchner, A., Rach, S., Lenarz, T., Dengler, R., Sandmann, P., 2017. Auditory and audio–visual processing in patients with cochlear, auditory brainstem, and auditory midbrain implants: an EEG study. Hum. Brain Mapp. 38 (4), 2206–2225.
Schierholz, I., Schönermark, C., Ruigendijk, E., Kral, A., Kopp, B., Büchner, A., 2021. An event-related brain potential study of auditory attention in cochlear implant users. Clin. Neurophysiol. 132(9), 2290–2305.
Schreitmüller, S., Frenken, M., Bentz, L., Ortmann, M., Walger, M., Meister, H., 2018. Validating a method to assess lipreading, audiovisual gain, and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head. Ear Hear. 39 (3), 503–516.
Senkowski, D., Pomper, U., Fitzner, I., Engel, A.K., Kral, A., 2014. Beta-band activity in auditory pathways reflects speech localization and recognition in bilateral cochlear implant users. Hum. Brain Mapp. 35 (7), 3107–3121.
Sharma, A., Dorman, M.F., Spahr, A.J., 2002. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. Ear Hear. 23 (6), 532–539.
Stevenson, R., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., Kurela, L. R., Siemann, J.K., James, T.W., Wallace, M.T., 2014. Identifying and quantifying multisensory integration: a tutorial review. Brain Topogr. 27 (6), 707–730.
Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2010. Does brain activity at rest reflect adaptive strategies? Evidence from speech processing after cochlear implantation. Cerebr. Cortex 20 (5), 1217–1222.
Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2013. Visual activity predicts auditory recovery from deafness after adult cochlear implantation. Brain 136 (12), 3682–3695.
Stropahl, M., Debener, S., 2017. Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. Neuroimage: Clinica 16, 514–523.
Stropahl, M., Plotz, K., Schönfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., Debener, S., 2015. Cross-modal reorganization in cochlear implant users: auditory cortex contributes to visual face processing. Neuroimage 121, 159–170.
Stropahl, M., Bauer, A.-K.R., Debener, S., Bleichner, M.G., 2018. Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. Front. Neurosci. 12, 309.
Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011.
Tremblay, C., Champoux, F., Lepore, F., Théoret, H., 2010. Audiovisual fusion and cochlear implant proficiency. Restor. Neurol. Neurosci. 28 (2), 283–291.
Ulrich, R., Miller, J., Schröter, H., 2007. Testing the race model inequality: an algorithm and computer programs. Behav. Res. Methods 39 (2), 291–302.
Vaughan Jr., H.G., 1982. The neural origins of human event-related potentials. Ann. N. Y. Acad. Sci. 388 (1), 125–138.
Viola, F.C., De Vos, M., Hine, J., Sandmann, P., Bleeck, S., Eyles, J., Debener, S., 2012. Semi-automatic attenuation of cochlear implant artifacts for the evaluation of late auditory evoked potentials. Hear. Res. 284 (1–2), 6–15.
Weglage, A., Müller, V., Layer, N., Abdellatif, K., Lang-Roth, R., Walger, M., Sandmann, P., 2022. Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users. Brain Topogr. 1–22.
Wesemann, W., Schiefer, U., Bach, M., 2010. Neue DIN-Normen zur Sehschärfebestimmung. Ophthalmologe 107 (9), 821–826.
Wesker, T., Meyer, B., Wagener, K., Anemüller, J., Mertins, A., Kollmeier, B., 2005. Oldenburg logatome speech corpus (OLLO) for speech recognition experiments with humans and machines. In: Ninth European Conference on Speech Communication and Technology.
Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data–a practical approach. J. Neurosci. Methods 250, 34–46.
Williams, N., 2017. The borg rating of perceived exertion (RPE) scale. Occup. Med. 67 (5), 404–405.
Williges, B., Wesarg, T., Jung, L., Geven, L.I., Radeloff, A., Jürgens, T., 2019. Spatial speech-in-noise performance in bimodal and single-sided deaf cochlear implant users. Trends. Hear. 23, 2331216519858311.
Woods, D.L., Knight, R., Scabini, D., 1993. Anatomical substrates of auditory selective attention: behavioral and electrophysiological effects of posterior association cortex lesions. Cognit. Brain Res. 1 (4), 227–240.
Zeng, F.-G., 2011. Advances in auditory prostheses. In: Auditory Prostheses. Springer, pp. 1–11.