



Distinct multisensory perceptual processes guide enhanced auditory recognition memory in older cochlear implant users

Jan-Ole Radecke^{a,b,c,*}, Irina Schierholz^{d,i}, Andrej Kral^{b,c}, Thomas Lenarz^{b,d},
Micah M. Murray^{e,f,g,h}, Pascale Sandmannⁱ

^a Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Germany

^b Institute of Audioneurotechnology, Hannover Medical School, Hannover, Germany

^c Department of Experimental Otolaryngology, ENT Clinics, Hannover Medical School, Hannover, Germany

^d Department of Otolaryngology, Hannover Medical School, Hannover, Germany

^e The LINE (The Laboratory for Investigative Neurophysiology), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

^f CIBM Center for Biomedical Imaging of Lausanne and Geneva, Lausanne, Switzerland

^g Department of Ophthalmology, Fondation Asile des aveugles, Lausanne, Switzerland

^h Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

ⁱ Department of Otorhinolaryngology, University of Cologne, Cologne, Germany

ARTICLE INFO

Keywords:

Cochlea implant
Audio-visual integration
Object recognition
Multisensory
Cross-modal
Rehabilitation
Congruency
Congruency Related Response

ABSTRACT

In naturalistic situations, sounds are often perceived in conjunction with matching visual impressions. For example, we *see* and *hear* the neighbor's dog barking in the garden. Still, there is a good chance that we recognize the neighbor's dog even when we only *hear* it barking, but do not *see* it behind the fence. Previous studies with normal-hearing (NH) listeners have shown that the audio-visual presentation of a perceptual object (like an animal) increases the probability to recognize this object later on, even if the repeated presentation of this object occurs in a purely auditory condition. In patients with a cochlear implant (CI), however, the electrical hearing of sounds is impoverished, and the ability to recognize perceptual objects in auditory conditions is significantly limited. It is currently not well understood whether CI users – as NH listeners – show a multisensory facilitation for auditory recognition. The present study used event-related potentials (ERPs) and a continuous recognition paradigm with auditory and audio-visual stimuli to test the prediction that CI users show a benefit from audio-visual perception. Indeed, the congruent audio-visual context resulted in an improved recognition ability of objects in an auditory-only condition, both in the NH listeners and the CI users. The ERPs revealed a group-specific pattern of voltage topographies and correlations between these ERP maps and the auditory recognition ability, indicating a different processing of congruent audio-visual stimuli in CI users when compared to NH listeners. Taken together, our results point to distinct cortical processing of naturalistic audio-visual objects in CI users and NH listeners, which however allows both groups to improve the recognition ability of these objects in a purely auditory context. Our findings are of relevance for future clinical research since audio-visual perception might also improve the auditory rehabilitation after cochlear implantation.

1. Introduction

The detection and recognition of complex natural sounds, in particular auditory objects (Murray et al., 2006; Shinn-Cunningham, 2008), is essential for the perception of the auditory environment. Following the predictive coding theory (Friston, 2005; Rao and Ballard, 1999; Rescorla

and Solomon, 1967), the recognition of auditory objects is likely based on recurrent and interactive bottom-up and top-down auditory processing. Specifically, the theory suggests that in response to an incoming auditory signal, previously established memory representations are recruited to build up top-down expectations about environmental signals and to guide the recognition of auditory scenes, objects or features

* Corresponding author at: Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20251 Hamburg, Germany.

E-mail address: janole.radecke@uni-luebeck.de (J.-O. Radecke).

<https://doi.org/10.1016/j.nicl.2022.102942>

Received 29 October 2021; Received in revised form 23 December 2021; Accepted 10 January 2022

Available online 12 January 2022

2213-1582/© 2022 The Authors.

Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(Bar, 2007; Gazzaley and Nobre, 2012; Kral et al., 2017; Näätänen et al., 2007; Rönnberg et al., 2013; Stefanics et al., 2014). Within this framework, incoming sensory information can be correctly classified, as either matching (hit) or mismatching (correct rejection) the previously formed expectation. On the other hand, if the sensory information is insufficient and cannot be mapped correctly onto prior expectations, this might provoke a mismatch (i.e. no clear match) and an additional explicit processing of the sensory information (compare Bar, 2007; Rönnberg et al., 2013). As a result, the object might be misclassified (miss, false positive), or might lead to correct classification (hit, correct rejection), while binding additional processing resources.

Perceptual expectations are influenced not only by intra-modal but also by cross-modal information. The recognition of complex naturalistic objects is even enhanced when these objects were previously experienced in a congruent multisensory context (Thelen and Murray, 2013). This indicates that a multisensory stimulus context allows more accurate access to memory representations which can be utilized for subsequent top-down unisensory stimulus recognition. Consistent with this view, it has been previously shown that an audio-visual context – when compared to a unisensory context – facilitates the later recognition of visual (Murray et al., 2005, 2004; Thelen et al., 2012) and auditory objects in normal-hearing (NH) listeners (Matusz et al., 2015; Thelen et al., 2015).

It is currently not well understood how multisensory facilitation in the context of auditory object recognition is affected by auditory deprivation and the use of hearing devices. In patients with severe to profound sensorineural hearing loss, cochlear implants (CI) can restore hearing by electrically stimulating the residual fibers of the auditory nerve (House and Urban, 1973; Simmons et al., 1965). However, electrical hearing is challenging due to the fact that the CI signal is characterized by a limited spectro-temporal specificity and a small dynamic range (Drennan and Rubinstein, 2008; Friesen et al., 2001; Kral, 2007; Wilson and Dorman, 2008). These limitations negatively affect auditory recognition and speech intelligibility (Lorenzi et al., 2006) and will likely lead to a mismatch between the (top-down) recruited auditory memory representations, and the incoming (bottom-up) electrical input in complex listening situations (compare Kral et al., 2017; and Rönnberg et al., 2013). This mismatch will be alleviated with increasing CI experience, but listening might remain effortful and may require additional or alternative neuronal resources (Finke et al., 2016, 2015).

Previous studies have shown that hearing loss and cochlear implantation can induce functional changes in the auditory cortex (Green et al., 2008; Klinke et al., 1999; Kral et al., 2006; Kral and Sharma, 2012; Sandmann et al., 2015; Sharma et al., 2002a, 2002b) and other cortical regions (Bottari et al., 2014; Chen et al., 2016; Heimler et al., 2014; review in Kral et al., 2016). These adaptations may also affect interactions between the different sensory modalities (Lomber et al., 2010; Rouger et al., 2007; Schierholz et al., 2015; Schorr et al., 2005; Strelnikov et al., 2015). In particular, post-lingually deaf CI users show cross-modal reorganization in the auditory cortex (Rouger et al., 2012; Sandmann et al., 2012) and reveal stronger audio-visual interactions in the auditory cortex when compared with normal-hearing listeners (Schierholz et al., 2017). These cortical alterations in CI users may allow better lip-reading abilities and enhanced audio-visual integration skills (Stropahl et al., 2017; Stropahl et al., 2015), fostering speech-comprehension recovery after cochlear implantation (Strelnikov et al., 2013). Furthermore, it has been recently suggested that audio-visual processing in postlingually deafened CI users – when compared to NH listeners – differs in terms of an increased attentional weighting of visual information in an audio-visual context (Butera et al., 2018), which might result in an enhanced benefit from (congruent) visual information. In sum, it can therefore be speculated that CI users show a strong facilitation in the recognition of unisensory auditory objects if these were preceded by congruent audio-visual stimuli.

The current study used electroencephalography (EEG) and a continuous recognition paradigm with a sequence of auditory and

audio-visual environmental objects (Matusz et al., 2015). As far as we are aware, this is the first EEG study to compare the effect of cross-modal context on later auditory recognition between CI users and NH listeners. We analyzed event-related potentials (ERPs) to auditory and audio-visual stimuli, reflecting earlier (P1-N1-P2-complex at 50–250 ms after auditory stimulus onset) and later cortical processing stages (N2 and P3a/P3b at 250–500 ms). Previous studies using *unisensory auditory stimuli* have reported that N1, N2 and P3a/P3b ERPs of CI users are reduced in amplitude and/or prolonged in latency when compared to NH listeners, suggesting difficulties in the sensory (N1) and the higher-level cognitive processing (N2 and P3a/P3b) of the limited CI input (Finke et al., 2016, 2015; Henkin et al., 2014; Henkin et al., 2009; Sandmann et al., 2009). ERP differences between CI users and NH listeners have also been reported in studies using basic *audio-visual stimuli*, pointing to an enhanced visual modulation of auditory ERPs in elderly CI users when compared to NH listeners (Schierholz et al., 2017; Stropahl et al., 2015). Based on these ERP results and on the observation of a multisensory facilitation effect in NH listeners, we predicted that implanted individuals show a significant audio-visual facilitation for auditory object recognition as well. Our behavioral results confirmed a facilitatory effect of the congruent audio-visual context on the later auditory object recognition, in both the CI users and the NH listeners. However, a group-specific pattern of ERP results suggests that the CI users' behavioral improvement is based on altered cortical processing of audio-visual objects, likely reflecting compensatory changes due to auditory deprivation and/or degraded sensory input after implantation.

2. Materials and methods

2.1. Participants and procedure

In total, twenty-seven participants, including twelve NH controls and fifteen post-lingually deafened CI users from the Hanover Medical School, were tested in the current study. Three CI users were excluded from the analysis, due to an undocumented usage of a hybrid CI device ($N = 1$) or due to neurologic confounds ($N = 2$). Therefore, twelve CI users (age: 60 ± 5 years (yr; mean \pm standard deviation (SD)); range 51–68 yr; 5 female) and twelve NH controls (age: 59 ± 5.5 yr; range 51–68 yr; 5 female), matched for handedness, gender, age and side of auditory stimulation, were included in the final sample. All CI users of the present sample were highly experienced (CI experience with the tested ear: 69 ± 29.7 months (mos); range 30–117 mos). Detailed individual information on the CI device, demographic factors and residual hearing are provided in Table 1. All participants had normal or corrected-to-normal visual acuity. Speech intelligibility in noise was assessed by the Göttinger sentence test (GöSa; Kollmeier and Wesselkamp, 1997). Auditory, visual (lip-reading) and audio-visual monosyllabic word recognition performance was evaluated (Schierholz et al., 2017; Stropahl et al., 2015) using words from the Freiburg Monosyllabic Word Test (Hahlbrock, 1970). Auditory stimuli were presented at 65 dB SPL. Participants were reimbursed for their participation. The study was conducted in agreement with the declaration of Helsinki and with the consent of the local ethical committee of the Hanover Medical School. Written informed consent was provided by every participant prior to the experiment.

2.2. Cognitive tests

Cognitive tests were applied to explore potential differences in cognitive abilities between CI users and NH listeners (Heydebrand et al., 2007; Mosnier et al., 2015). Working memory was assessed by the verbal digit and the non-verbal block span tests (forward and reverse order) from the Wechsler Memory Scale (WMS-R; Härting et al., 2000). Furthermore, non-linguistic immediate and delayed open-set recall, as well as recognition were tested, in order to exclude general memory deficits (Rey Complex Figure Test (RCFT); Meyers and Meyers, 1995).

Table 1

Detailed information on CI users included in the study. Gender is indicated as F (female) or M (male). Implant devices that were used for stimulation during the experiment are marked with bold text. “Duration of Deafness” and “CI Experience” were quantified with respect to the ear that was stimulated during the experiment. Word Recognition was measured using the Freiburg monosyllabic word test (Hahlbrock, 1970). Göttinger sentence test represents the hearing in noise ability (n.a. = not applicable, due to technical reasons). 4PTA represents the pure tone average for four frequencies (0.5, 1, 2, 4 kHz).

#	Gender	Age [y]	Etiology	Hearing Device (Right/Left)	Duration of Deafness [m]	CI Experience [m]	Word Recognition (Stimulation Side) [%]	Göttinger Sentence Test [SNR]	4PTA (Right/Left) [dB HL]
1	F	53	unknown	AB HiRes90K Helix / AB HiRes90K Helix	3	81	45	11.8	> 80 / > 80
2	M	63	unknown	MedEl Concerto Standard /Hearing Aid	563	57	95	6.5	> 80 / 66
3	M	64	unknown	AB HiRes90K Advantage /AB HiRes90K Advantage	9	30	95	8.8	> 80 / > 80
4	F	51	acute hearing-loss	Nucleus CI422 / -	8	40	85	10.6	73 / 19
5	F	59	unknown	- / Nucleus CI512	6	64	90	2.1	> 80 / > 80
6	M	65	cholesteatoma	Nucleus CI512 / -	155	54	100	-1.9	73 / 10
7	M	58	unknown	AB HiRes90K Helix / AB HiRes90K Helix	36	106	75	4.1	> 80 / > 80
8	F	60	genetic	AB HiRes90K Helix/ AB HiRes90K Helix	1	52	80	7.7	> 80 / > 80
9	M	68	acute hearing-loss	Nucleus CI24RE / Hearing Aid	111	117	85	3.9	> 80 / 54
10	F	60	meningitis	Nucleus CI512 /Nucleus CI24RE	614	72	30	n.a.	> 80 / > 80
11	M	56	unknown	- / MedEl Concerto Flex EAS28*	1	42	55	9.7	> 80 / > 80
12	M	57	genetic	Nucleus CI 24R /Nucleus CI24RE	73	115	15	17.2	> 80 / > 80

The MWT-B (Mehrwachwahl-Wortschatz-Intelligenz-Test; Lehl, 1999) was applied to assess lexical recognition. Age-appropriate percentile ranks were evaluated for each test and respective subtests. The obtained values were compared between groups by means of separate independent *t*-tests.

2.3. Continuous recognition task

The experimental paradigm was based on a version of the continuous recognition paradigm described in Matusz et al. (2015; Fig. 1). In this paradigm, participants were presented with a continuous sequence of auditory and audio-visual environmental objects. The participants performed a two-alternative forced choice task, in which they pressed a key on a keyboard to indicate whether the current stimulus occurred for the first time (initial presentation) or whether it was previously presented in the current experimental block (repetition trial). The first presentation

of each object was either auditory (stimulus condition A) or audio-visual. For the audio-visual stimulus conditions, the sound object was either paired with a congruent pictogram (stimulus condition AVc) or with a meaningless picture (stimulus condition AVm). The repetition of an object was always purely auditory, here referred to as the stimulus conditions A- (following stimulus condition A), A+c (following stimulus condition AVc), and A+m (following stimulus condition AVm; Fig. 1A).

In total, 54 objects were presented twice (initial presentation, repetition trial) in each of six experimental blocks. Each block consisted of 108 trials, resulting in overall 648 trials across the experiment. The interval between the first occurrence of an object and its repetition was pseudo-randomly set to three, six or nine intervening trials. The order of trials within blocks and the order of blocks were counter-balanced to avoid sequence effects.

Participants performed a training sequence in order to familiarize with the task. In addition, a closed-set identification block was

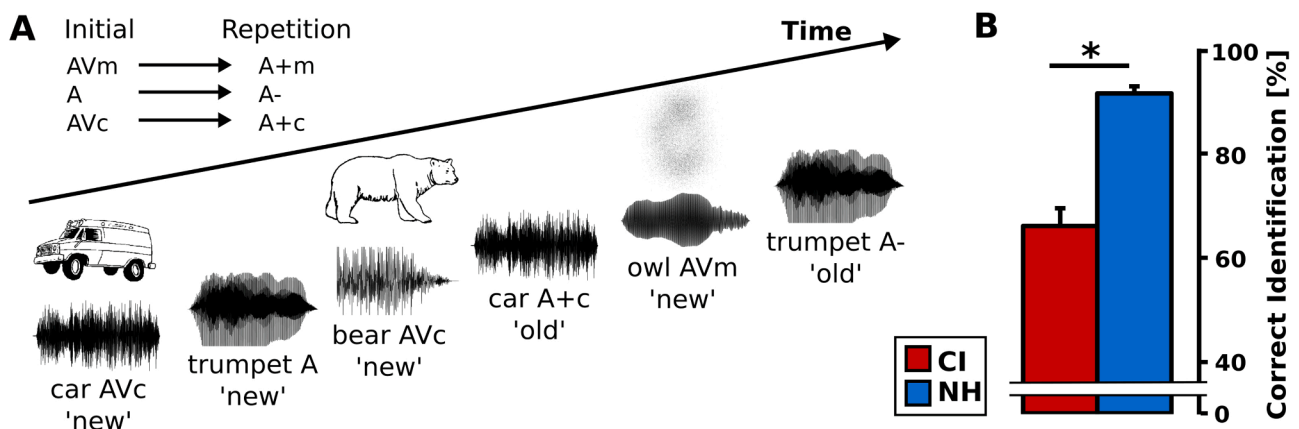


Fig. 1. Continuous recognition paradigm. A) During a continuous recognition paradigm, naturalistic auditory objects were presented twice during each of six experimental blocks. Initial auditory stimuli were either unisensory (A) or they were paired either with a meaningless visual stimulus (AVm) or paired with a congruent visual stimulus (AVc). Repetitions of objects were always auditory only (A+m, A-, A+c, respectively). B) A closed-set identification block revealed that CI users can identify the naturalistic auditory stimuli well above chance-level of 33%. Nevertheless, NH participants showed higher identification scores when compared with CI users. Mean ± standard error of the mean is presented, respectively.

conducted prior to the continuous recognition paradigm to validate the auditory intelligibility of the environmental objects (Fig. 1B). Herein, participants had to match each auditory object from the experiment to one of three pictograms, presented on the screen via button press.

2.4. Audio-visual stimuli

The sound stimuli and congruent pictograms of the environmental objects were taken from a set of audio-visual stimuli, applied in previous studies (e.g. Matusz et al., 2015). Meaningless versions of the pictograms were computed as scrambled versions of the meaningful pictogram's pixels. Sounds were linearly ramped for 10 ms at the onset and the offset to avoid click perceptions. Auditory stimulation was provided to one ear only. For CI users, the RMS-normalized sounds were presented via free-field from two loudspeakers (HECO victa 301), each located at 26° azimuth to the left and right side of the participant. For bilateral CI users, the subjectively reported better ear was selected for stimulation. Hearing devices of the contralateral ear (CI, hearing aid) were detached for the time of the experiment. The respective ear was additionally closed with an earplug. NH controls were stimulated on the side corresponding to their matched CI counterpart by means of an insert earphone (3 M E-A-RTONE 3A). The non-stimulated ear, however, was provided with a silent insert earphone. Sound volume was set to 65 dB SPL. As used in previous studies (Sandmann et al., 2015, 2010, 2009; Schierholz et al., 2017; Stropahl et al., 2015), a seven-point subjective loudness scale was applied to ensure sufficient and comfortable perceptual volume for each participant.

Congruent and meaningless pictures consisted of black pictograms on a white background with a Michelson contrast of 85% (background: 327 cd/m²; picture: 27 cd/m²; Konica Minolta LS 100). Pictures spanned about 7° of visual angle (congruent: 6.6° ± 0.6°, meaningless: 7.4° ± 0.7°). The pure auditory stimulus condition (A) was accompanied by a white background (327 cd/m²), which was also displayed after each trial. Visual stimuli were presented on a 27-inch screen (1920 × 1080 × 32 bit, 60 Hz refresh rate) at 1.5 m distance.

Auditory stimuli lasted for 500 ms. In the audio-visual stimulus conditions (AVc, AVm), pictures were presented with a stimulus onset asynchrony (SOA) of -130 ms visual lead. This SOA was chosen to approximate the time window of multisensory integration, which has been previously reported for audio-visual speech (van Wassenhove et al., 2007), while ensuring that the visual stimulus still affects auditory processing (semantic priming). Offset of pictures and sounds were synchronous. During each trial, a fixation point was presented at the center of the screen, 300 ms before stimulus onset. The post-stimulus interval was jittered uniformly between 1700 and 2100 ms in steps of 100 ms. The software Presentation (Neurobehavioral Systems, Berkeley, CA, USA) was employed for stimulus presentation and documentation.

2.5. Behavioral data analysis

We computed hit rates (% of correctly identified repetition trials), reaction times (of correct responses), as well as sensitivity and bias based on signal detection theory (i.e. d' and c , respectively; Macmillan and Creelman, 2005; see also Matusz et al., 2015). Sensitivity indicates the perceptual discriminability between each pair of initial and repeated object presentations (AVc and A+c, AVm and A+m, A and A-). It was computed separately for each pair of stimulus conditions, using probabilities of correct responses (HIT) and false positive responses (FP) as $d' = [z(\text{HIT}) - z(\text{FP})]$. The function $z(p)$, $p \in [0, 1]$, represents the inverse of the cumulative distribution function of the standard normal distribution. To control for possible extreme values, HIT and FP were computed as $p(x) = \frac{x+0.5}{N+1}$, where x reflects the sum of correct responses or false positives, respectively, and N reflects the corresponding number of trials. Correct responses were defined as the first correct response to a repetition trial, occurring between 100 ms post-stimulus and the onset of

the next trial. First presentations that were mistakenly identified as repetitions were labeled as false positive response. d' values will be denoted as recognition conditions d'_{AVm} , d'_A and d'_{AVc} , for the trials with meaningless audio-visual, pure auditory and congruent audio-visual stimuli, respectively. Furthermore, response bias c (c_{AVm} , c_A , c_{AVc}) was computed as $c = [0.5 * z(\text{HIT}) + z(\text{FP})]$. Sensitivity (d'_{AVm} , d'_A and d'_{AVc}), response bias (c_{AVm} , c_A , c_{AVc}) and repetition hit rates (% for A+m, A-, A+c) were evaluated using separate mixed model repeated-measures ANOVAs with Recognition Condition (sensitivity: d'_{AVm} , d'_A , d'_{AVc} ; response bias: c_{AVm} , c_A , c_{AVc} ; hit-rates: % for A+m, A-, A+c) as the within-subject factor and Group (CI, NH) as the between-subject factor. Median reaction times (RT) of correct responses were computed for every block and averaged for each participant. RTs were evaluated separately for initial presentations and repetitions by means of two mixed model repeated-measures ANOVAs with Stimulus Condition (initial: AVm, A, AVc; repetition: A+m, A-, A+c) as the within-subject factor and Group (CI, NH) as the between-subject factor. Follow-up t -tests were computed and corrected for multiple comparisons using Bonferroni-correction, respectively.

Nonparametric Spearman rank correlations were computed between behavioral parameters from the experiment on the one hand, and demographic variables and cognitive test measures on the other hand. The p -values were corrected using the Bonferroni-Holm correction (Holm, 1979).

2.6. EEG data acquisition and preprocessing

EEG data were recorded using a BrainAmp EEG amplifier system (BrainProducts, Gilching, Germany) and a sampling rate of 1 kHz. An analog filter between 0.02 and 250 Hz was applied. Ninety-six Ag/AgCl electrodes were placed in an equidistant layout, with the online-reference placed at the nose tip and a fronto-polar ground electrode (Easycap, Herrsching, Germany). Two electrodes were placed below the eyes to record the electrooculogram. Electrode impedances were kept below 20 kΩ. Electrodes around the CI processor and the transmitting coil were spared and excluded from the data analysis. Among the CI users, 7.6 ± 0.1 electrodes were omitted this way.

EEG data were processed using custom MATLAB scripts (The Mathworks Ltd., Natick, MA, USA) and the EEGLAB toolbox (Delorme and Makeig, 2004). Data were digitally down-sampled to 500 Hz. To correct for typical electrophysiological artifacts, an extended infomax independent component analysis (ICA) was applied to the data (Bell and Sejnowski, 1995; Lee et al., 1999). For this purpose, continuous data were high- and lowpass-filtered between 1 and 40 Hz and cut into epochs of 1 s duration. Artefactual components related to vertical eye-movements, eye-blinks, heart-beat, muscle activity and the electrical artifact from the CI were identified by visual inspection (Debener et al., 2008; Sandmann et al., 2009; Viola et al., 2011; Viola et al., 2010). ICA components which reflect the electrical artifacts in the EEG of CI users show a pedestal in the time course and a centroid in the topography which is lateralized to the side of stimulation (Sandmann et al., 2009; Viola et al., 2011). The weights of identified artifact components were set to zero and applied to the raw data. ICA-corrected data were digitally filtered between 0.4 and 40 Hz half amplitude cutoff, using separate Kaiser-windowed zero-phase FIR-filters (Kaiser-β = 5.65) for high-(transition bandwidth: 2 Hz) and low-pass filtering (transition bandwidth: 0.8 Hz; Widmann et al., 2014). Epochs were extracted from -700 to 1000 ms, centered on the auditory stimulus onset. Data were corrected using a 200 ms baseline before fixation onset, respectively. Non-stereotyped artefactual epochs were rejected using a joint probability criterion of 3 SD, resulting in 74 trials for every subject and stimulus condition on average. Finally, data were re-referenced to a common average reference, and channels missing due to the CI device were interpolated using a spherical spline. In general, the signal-to-noise ratios of the resulting ERPs were evaluated and did not differ between groups (see Supplementary Material A).

2.7. Analysis of Event-Related potentials (ERPs)

ERPs from correct trials were computed separately for CI users and NH listeners for both the initial presentations (AVm, A, AVc) and the repetition trials (A+m, A-, A+c). As detailed below, our ERP analyses mainly focused on the two audio-visual conditions from *initial* presentations in order to explore the electrophysiological correlates of the multisensory congruency effect on recognition memory (AVc vs. AVm). Comparing the ERPs between these two conditions (AVc vs. AVm) allows to draw inferences about the specific effect of additional *congruent* visual information on cortical multisensory processing as well as the associated improvement in auditory recognition (reviewed in [Matusz et al., 2017](#)). Moreover, prior research has shown that modulations in brain responses to multisensory stimuli – but not to unisensory stimuli – are correlated with memory performance ([Thelen et al., 2014](#)).

2.7.1. ERP analysis: Global field power (GFP) and global map dissimilarity (GMD)

For each group, difference ERPs were calculated for initial object presentations (CI_{AVc - AVm} and NH_{AVc - AVm}) and their repetitions (CI_{A+c - A+m} and NH_{A+c - A+m}). These difference ERPs were then separately compared between the two groups. We evaluated these ERP group differences (CI_{AVc - AVm} vs. NH_{AVc - AVm} and CI_{A+c - A+m} vs. NH_{A+c - A+m}) by computing the *global field power* (GFP) and the *global map dissimilarity* (GMD) to quantify differences in response strength and response topography, respectively ([Murray et al., 2008](#)).

The GFP was computed as $GFP_x = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$ and $GFP_y = \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2}$, respectively. Herein, n represents the number of electrodes and x_i and y_i represent the two ERP signals at the i -th electrode, respectively. The GMD, by contrast, is a measure to identify differences in ERP topographies (and by extension the underlying neural generator configurations) between two experimental conditions, independent of signal strength ([Lehmann and Skrandies, 1980](#); [Murray et al., 2008](#)). It is computed as $GMD(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i}{GFP_x} - \frac{y_i}{GFP_y} \right)^2}$. Herein, n represents the number of electrodes and x_i and y_i represent the two average-referenced ERP signals (CI_{AVc - AVm} and NH_{AVc - AVm} or CI_{A+c - A+m} and NH_{A+c - A+m}, respectively) at the i -th electrode, respectively. In other words, the GMD computes the sample-by-sample Euclidean distance between reference-independent spatio-temporal field distributions across the scalp with n electrodes of two ERP signals x and y , normalized by the respective GFP.

Group differences in GFP (ERP strength) and GMD (ERP topography) were analyzed by means of a non-parametric permutation test ([Maris and Oostenveld, 2007](#)) computed at every sample point of the ERP segment. The permutation test compared the ERP amplitudes between the two groups (GFP) and explored for a significant spatial correlation between the ERP maps of the two groups (GMD). These analyses were conducted on the difference ERPs separately for the initial trials (CI_{AVc - AVm} vs. NH_{AVc - AVm}) and the repetition trials (CI_{A+c - A+m} vs. NH_{A+c - A+m}) by using 5000 permutations and an FDR-correction across the time samples to control for multiple comparisons (*false discovery rate*; [Benjamini and Hochberg, 1995](#)).

Unisensory auditory ERPs (A, A-) were not included in the GFP and GMD analysis, due to the nature of auditory ERPs which is inherently different from audio-visual ERPs (compare [Fig. 3](#) and [Supplementary Material B](#)).

2.7.2. Topographic clustering analyses

To further explore the ERP differences between CI users and NH listeners, the group-averaged data from the AVc and AVm conditions from both groups were submitted to a hierarchical clustering using the CARTOOL freeware ([Brunet et al., 2011](#)), identifying template topographies in circumscribed time windows of interest. Specifically, we used the atomize and agglomerate hierarchical cluster analysis that has been

designed for ERPs and considers the global explained variance of a cluster and thus prevents blindly combining or agglomerating clusters of relatively short duration. Conceptually, topographic clustering identifies the minimal set of topographies that account for (i.e. explain the greatest variance) a given dataset (here: the pooled group-averaged ERPs from both conditions and groups). This is predicated on the empirical observation that the ERP topography does not vary randomly across time, but rather remains in a stable configuration for periods of time before assuming a next, different topography. This has been commonly referred to as microstates (reviewed in [Michel and Koenig, 2018](#)). A full description of the approach can be found in [Brunet et al. \(2011\)](#) as well as ([Murray et al., 2009](#)).

Single-subject fitting of the identified template topographies ([Murray et al., 2008](#)) was performed to evaluate the preponderance of specific voltage topographies during the AVm and AVc conditions on the single-subject level. More specifically, for each subject and condition, sample-wise spatial correlations were computed between the observed voltage topographies and each template topography. Each sample was assigned to the one template map showing the highest spatial correlation. As output, we obtained the total number of samples that were assigned to each template topography.

These total time samples were evaluated using mixed model repeated-measures ANOVAs with initial Stimulus Condition (AVm, AVc) and Template Map as the within-subject factors and Group (CI, NH) as the between-subject factors, separately for each time-window of interest. Group-wise mixed model repeated-measures ANOVAs (Condition \times Template Map) were computed in case of significant three-way interactions. Finally, follow-up t -tests were computed and corrected for multiple comparisons using the Bonferroni-Holm correction ([Holm, 1979](#)), respectively. Correlations were computed between the discriminabilities (d'_{AVc} , d'_{AVm}) and the preponderance of condition-specific topographies during the respective condition (AVc, AVm), separately for each group. Non-parametric Spearman rank correlation coefficients (ρ) were computed and p -values were corrected using the Bonferroni-Holm correction ([Holm, 1979](#)).

2.8. Statistical analysis software

IBM SPSS Statistics (IBM Corp., Armonk, NY, USA) and MATLAB were utilized for statistical analyses. In general, for both behavioral and electrophysiological data, significance levels were set to $\alpha = 0.05$, and a Greenhouse-Geisser correction was applied in case the sphericity assumption was violated. If not indicated otherwise, mean (M) and standard error of the mean (SEM), as well as corrected p -values are reported. Partial η^2 (η_p^2) and r were denoted as estimates of effect size for parametric tests, respectively.

3. Results

3.1. Behavioral data

3.1.1. Discrimination ability, recognition and reaction times

In the closed-set identification block, task performance was above chance level in both groups (CI users: $67.4\% \pm 3.9\%$; NH listeners: $91.5\% \pm 1.5\%$), indicating that both groups were able to correctly recognize the sound objects ([Fig. 1B](#)). NH listeners showed an enhanced ability to identify auditory objects when compared with CI users ($t_{14,2} = -7.2$, $p < .001$, $r = 0.86$).

Regarding the continuous recognition paradigm, both groups showed enhanced *hit rates* for auditory objects if they were presented previously in a congruent audio-visual context (A+c) as compared to an audio-visual meaningless (A+m) or an auditory-only stimulus context (A-; [Table 2](#)). For the hit rates, a mixed-model repeated-measures ANOVA revealed main effects for Stimulus Condition ($F_{2,44} = 27.94$, $p < .001$, $\eta_p^2 = 0.56$) and Group ($F_{1,22} = 5.63$, $p = .027$, $\eta_p^2 = 0.20$). However,

Table 2

Mean hit rates and SEM for auditory repetition stimulus conditions, following initial audio-visual meaningless pairing (A + m), auditory objects (A-), or audio-visual congruent pairing (A + c).

Hit-Rates [%]	CI	NH
A + m	72.2 ± 3.8	81.6 ± 2.6
A-	73.9 ± 3.7	83.1 ± 2.4
A + c	80.6 ± 3.0	90.4 ± 1.7

the Group × Stimulus Condition interaction was not significant ($p = .941$). Hit rates for stimulus condition A + c were significantly increased when compared to A + m ($t_{23} = 6.95, p < .001, r = 0.82$) and A- ($t_{23} = 5.62, p < .001, r = 0.76$). However, no difference was observed between hit rates for A + m and A- ($t_{23} = -1.39, p = .53$). In sum, the results of the hit rates indicate that both groups have a multisensory facilitation on the subsequent recognition of auditory objects specifically in the congruent audio-visual recognition condition.

Next, we conducted analyses of d' and c to ascertain if the above multisensory facilitation was more likely due to perceptual sensitivity or response bias. A Group (CI, NH) × Recognition Condition (AVm, A, AVc) repeated-measures ANOVA for d' revealed main effects of Group ($F_{1, 22} = 27.1, p < .001, \eta_p^2 = 0.6$) and Recognition Condition ($F_{1.5, 32.2} = 109.3, p < .001, \eta_p^2 = 0.83$). The Group × Recognition Condition interaction was not significant for the d' values ($p = .462$). Overall, CI users showed a significantly reduced d' when compared to NH participants. Furthermore, post-hoc analysis revealed a significantly higher d' in both groups (Fig. 2A) when the initial object had been a congruent audio-visual stimulus pair, compared to when it was either an audio-visual meaningless pair or a unimodal auditory stimulus ($d'_{AVc} > d'_{AVm}; t_{23} = 11.2, p < .001, r = 0.92; d'_{AVc} > d'_A; t_{23} = 13.5, p < .001, r = 0.94$). Moreover, higher sensitivities (Fig. 2A) were also observed for initial unisensory auditory compared to audio-visual meaningless objects ($d'_{AVm} < d'_A; t_{23} = -3.2, p < .004, r = 0.56$). Importantly, the response bias c was comparable across all recognition conditions (effects of Group, Recognition Condition, Group × Recognition Condition: all $p > .05$; Fig. 2B).

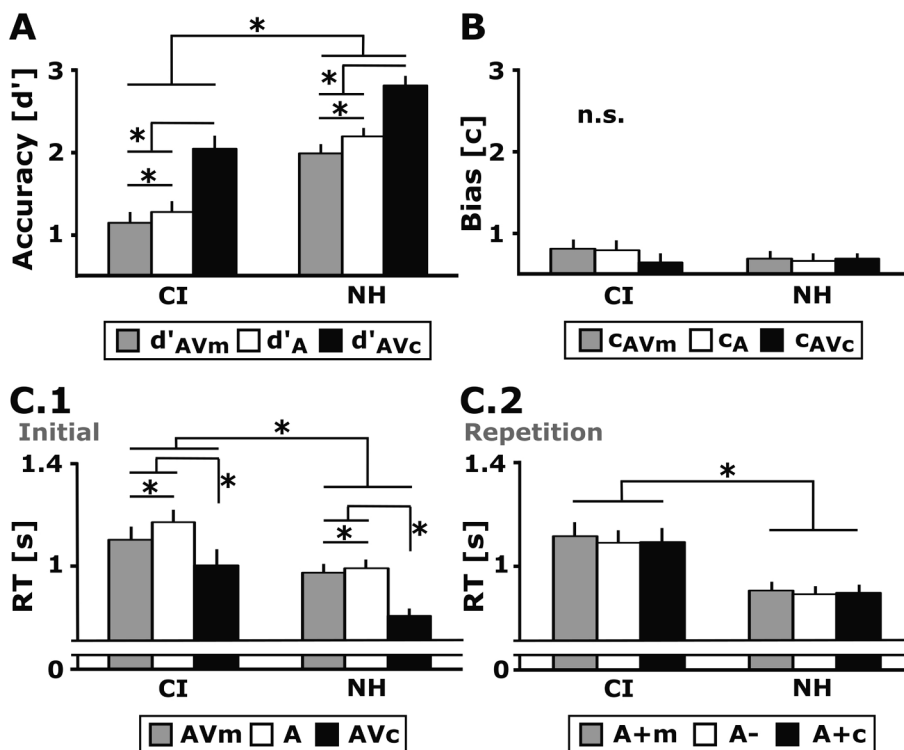


Fig. 2. Behavioral results of the CI users and the NH listeners. A) Mean and standard error of the mean (SEM) are given for the sensitivity index d' , B) the response bias c , and C) the RTs for the initial (C.1) and repetition (C.2) stimulus conditions. The results indicate increased d' values in the audio-visual congruent recognition condition in both the CI users and the NH listeners (A). By contrast, no significant differences between conditions or groups were observed for the response bias (B). For initial stimulus presentations, RTs were faster for the audio-visual congruent compared to the audio-visual meaningless and the auditory-only stimulus condition (C.1). By contrast, repetition trials revealed no stimulus-condition effect for RTs (C.2). Overall, NH participants showed higher accuracies and faster reaction times compared to CI users.

Median RTs for *initial stimulus presentations* indicated overall slowed RTs for CI users when compared to NH listeners (Main effect of Group: NH < CI: $F_{1, 22} = 8.6, p = .008, \eta_p^2 = 0.28$). Further, we observed a main effect of initial Stimulus Condition ($F_{1.4, 30.4} = 31.63, p < .001, \eta_p^2 = 0.59$), which was caused by faster RTs in stimulus condition AVc, compared to stimulus condition AVm, which in turn was significantly faster when compared to stimulus condition A ($RT_{AVc} < RT_{AVm}; t_{23} = -5.1, p < .001, r = 0.73; RT_{AVc} < RT_A; t_{23} = -6.6, p < .001, r = 0.81; RT_{AVm} < RT_A; t_{23} = -3, p = .006, r = 0.53$; Fig. 2C.1). There was no Group × Stimulus Condition interaction ($p = .320$). For the *repetition trials*, a significant main effect of Group was observed, indicating overall slowed RTs in CI users compared to NH listeners (NH < CI: $F_{1, 22} = 11.7, p = .002, \eta_p^2 = 0.35$; Fig. 2C.2). However, there was no main effect of repetition Stimulus Condition ($p = .229$), nor a Group × Stimulus Condition interaction ($p = .818$).

In sum, these behavioral results indicate for both the CI users and the NH listeners a multisensory facilitation of the recognition of auditory objects specifically by preceding *congruent* audio-visual stimuli (AVc).

3.1.2. Speech recognition ability and cognitive tests

Speech recognition ability of monosyllabic words (presented without background noise) was high in both the NH listeners (A and AV: 100%) and the CI users (A: $71 \pm 8.1\%$, AV: $89 \pm 2.1\%$; AV > A: $t_{11} = 2.52, p = .028, r = 0.60$). Lip-reading skills were comparable between the two groups (NH: $11 \pm 2.3\%$, CI: $19 \pm 4.1\%$; $p = .132$). Speech-in-noise intelligibility – assessed by the GöSA sentence test – was significantly worse in CI users compared to NH listeners (NH: -4.2 ± 0.2 , CI: 7.3 ± 1.4 ; NH < CI: $t_{10.4} = 7.29, p < .001, r = 0.91$).

Group-specific percentile ranks of the cognitive tests are provided in Table 3. Regarding the tests assessing the working memory capacity, the percentile ranks did not significantly differ between the CI users and the NH listeners, neither for the verbal digit span task ($p = .119$) nor for the non-verbal block span task ($p = .519$). Furthermore, general memory function, as assessed by the non-verbal Rey Complex Figure Test (RCFT), showed comparable performance for the immediate ($p = .982$) and the delayed free recall ($p = .788$), as well as for the cued recognition ($p =$

Table 3

Mean percentile ranks and SEM for the applied cognitive tests are shown together with the *t*-values derived from the between-group comparisons. Results are provided for verbal and non-verbal working memory tasks (verbal digit span and non-verbal block span). Furthermore, non-linguistic memory functions were validated using the Rey Complex Figure Test (RCFT) for immediate and delayed recall as well as cued recognition of an abstract pictogram. As indicated by the asterisk, lexical recognition (MWT-B) was significantly decreased in CI users when compared to NH participants (*p* = .003). Details on statistical results are provided in the text.

Cognitive Test	CI	NH	<i>t_{df}</i>	<i>df</i>
Digit Span (WMS-R)	34.6 ± 9.0	55.7 ± 9.4	-1.62n.s.	22
Block Span (WMS-R)	69.0 ± 9.5	60.8 ± 8.1	0.66n.s.	22
Immediate Recall (RCFT)	72.1 ± 8.8	72.4 ± 6.3	-0.23n.s.	22
Delayed Recall (RCFT)	68.4 ± 9.1	71.4 ± 6.1	-0.27n.s.	19.3
Recognition (RCFT)	59.0 ± 9.8	71.9 ± 7.4	-1.17n.s.	22
Lexical Recognition (MWT-B)	59.3 ± 6.5	85.0 ± 3.8	-3.39 *	22

.254) of the complex figure. These results indicate that no general memory deficits were observed in CI users, compared to NH listeners.

Finally, lexical access, as assessed by the MWT-B percentile ranks, was significantly lower for CI users compared to NH controls (NH > CI: *t*₂₂ = -3.39, *p* = .003, *r* = 0.59).

3.2. Event-related potentials

3.2.1. ERPs to repetition trials

Auditory ERPs, evoked by the physically identical repetition stimulus conditions (A + m, A-, A + c), showed typical ERP components, reflecting earlier (P1-N1-P2 complex at 50–250 ms after auditory stimulus onset) and later cortical processing stages (anterior N2 and P3a around 250–500 ms; Supplementary Fig. 2). Evaluation of the signal-to-noise ratio of the ERPs indicated a sufficient correction of the electrical artifact, an appropriate restoration of ERPs, and no group differences in

the signal-to-noise ratio despite unequal trial numbers across the two groups (see Supplementary Material A). However, auditory repetition ERPs were similar across stimulus conditions and no group-specific differences in GMD or GFP were observed (CI_{A+c} - A+m vs. NH_{A+c} - A+m; Supplementary Material B). Thus, we did not observe neurophysiological evidence for the modulation of congruent audio-visual stimuli on subsequent auditory processing, as was indicated by the modulation of recognition hit-rates following the AVc condition (section 3.1.1, Table 2) (compare Matusz et al., 2015).

3.2.2. ERPs to initial trials

In contrast to the repetition stimulus conditions, the ERPs from the initial trials showed remarkable differences between the audio-visual congruent (AVc) and the audio-visual meaningless (AVm) stimulus conditions in both groups (Figs. 3, 4 and 5). Descriptively, both groups showed an early visual P1 ERP component (vP1) over posterior scalp regions at approximately 80 ms after visual stimulus onset (Fig. 3C and F). We also registered audio-visual P1 and N1 responses (with respect to auditory stimulus onset) at central sites across both groups and stimulus conditions (Fig. 3B and E). With regards to frontal scalp regions (Fig. 3A and D), the difference wave between the two audio-visual stimulus conditions (AVc - AVm) revealed a positive deflection in both groups, peaking at around 200 ms, in the following referred to as *Congruency-Related Response* (CRR). Descriptively, the CI users' CRR response (Fig. 3A) appeared to be extended in time (CI: 80 to 430 ms; NH: 140 to 330 ms) and of larger amplitude (CI: 2.5 μV; NH: 1.5 μV) when compared with NH listeners (Fig. 3D). Finally, we observed a P3b component over parieto-occipital scalp regions in the group of NH listeners, which was particularly pronounced for the congruent audio-visual stimulus condition (Fig. 3F) and which peaked around 350 ms (2.5 μV) after auditory stimulus onset.

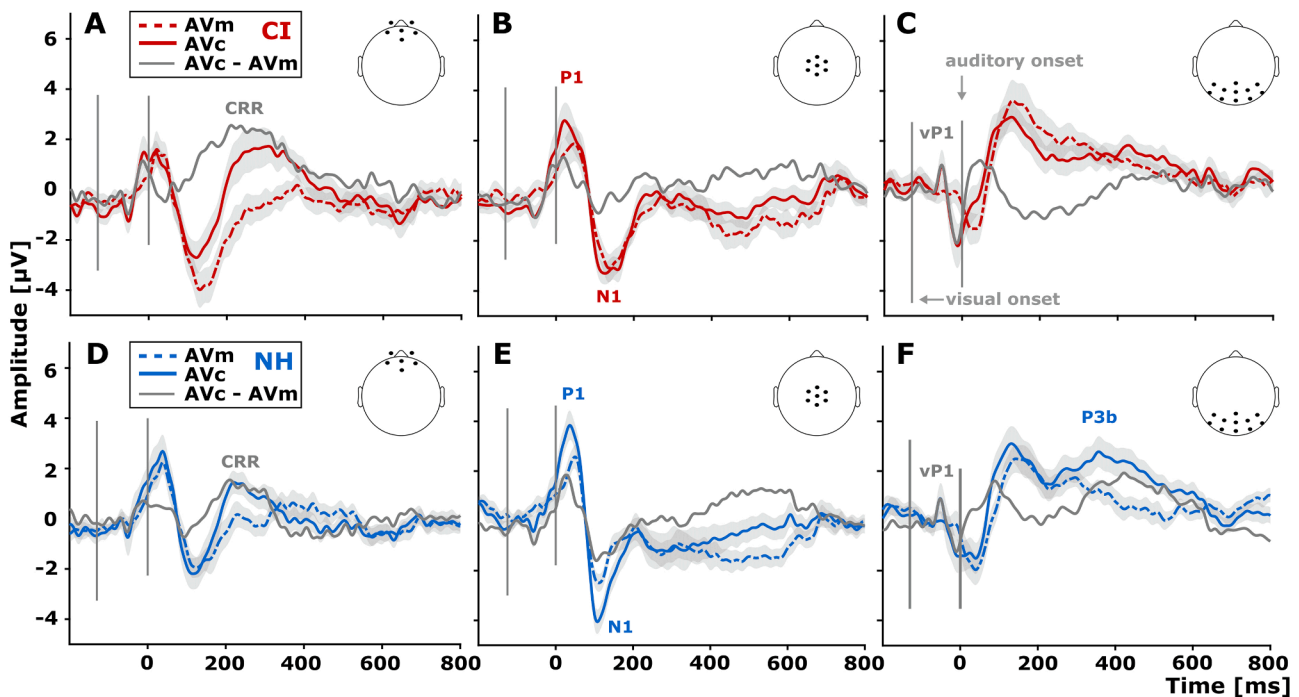
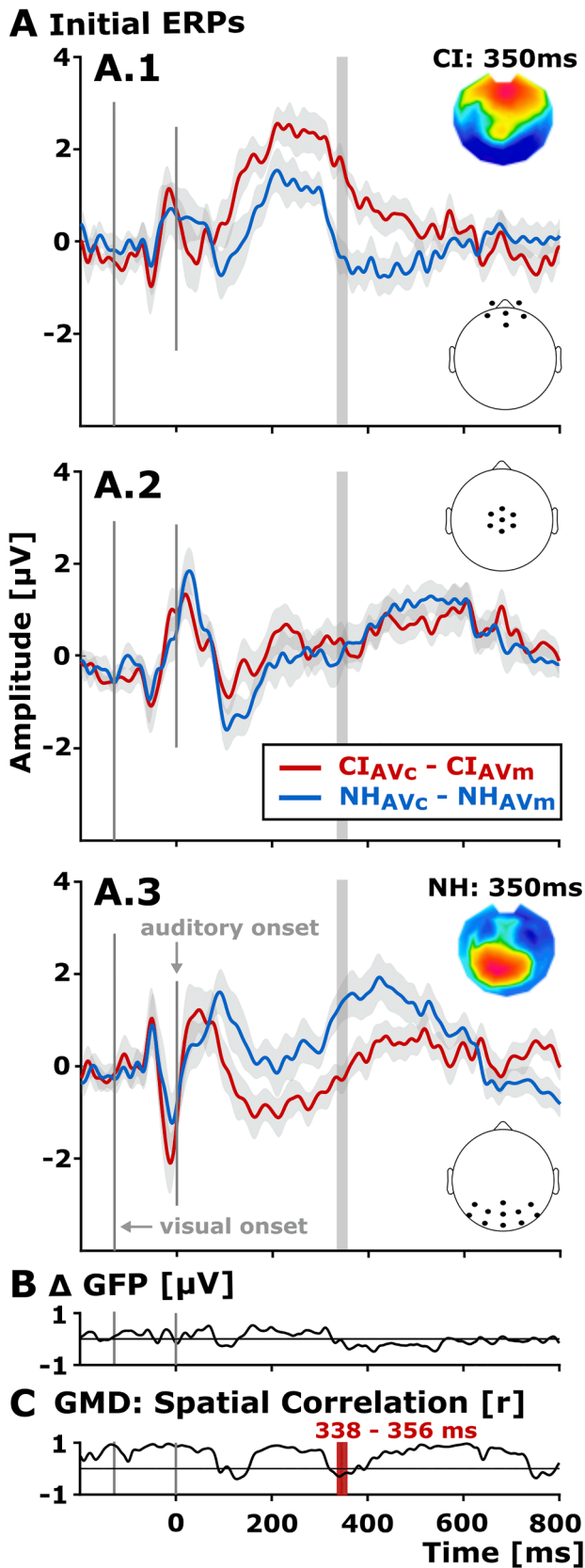


Fig. 3. ERPs to initial audio-visual stimulus conditions (AVm, AVc) for CI users (A-C) and NH participants (D-F). ERPs are shown for a frontal (A, D), a central (B, E) and a parieto-occipital (C, F) electrode cluster. The central electrode cluster (B, E) reveals an audio-visual P1 and N1 component in both groups. Further, the difference wave (AVc - AVm) shows a positivity in the frontal electrode cluster in both groups (A, D), which is inverted in the posterior electrode cluster (C, F). Note also the P3b component in response to the congruent audio-visual stimulus condition over posterior scalp locations, particularly pronounced in the group of NH participants (F). Mean ERPs for the respective stimulus conditions are represented by the colored dashed and solid lines. Grey patches indicate the SEM for each ERP. Solid grey lines represent the difference ERP (CI_{AVc - AVm} and NH_{AVc - AVm}). Electrode clusters are indicated on a model head, viewed from top.



(caption on next column)

Fig. 4. Comparison of the difference waves between the CI users and the NH listeners. A) Difference ERPs for CI users (red) and NH participants (blue) for *initial* audio-visual stimulus conditions (CI_{AVc} - AV_m and NH_{AVc} - AV_m, A.1–3) for three different electrode clusters (from top to bottom: frontal, central, parieto-occipital). Grey patches indicate the SEM for each difference ERP. Electrode clusters are indicated on a model head, viewed from top. Note the ERP maps of the differences waves (CI_{AVc} - AV_m and NH_{AVc} - AV_m) at 350 ms, showing a frontal positivity in CI users – referred to as the CRR response (A.1) – and a posterior positivity in NH listeners – referred to as the P3b response (A.3). B) Comparison of the global field power between the CI users and NH listeners. The figure illustrates the difference of GFPs (Δ GFP) between groups and conditions (CI_{AVc} - AV_m - NH_{AVc} - AV_m). There was no significant group difference in the ERP response strength at any sample point of the ERP segment. C). Topographic group differences were analyzed by the GMD and illustrated here by the sample-wise spatial correlations between the difference ERP topographies of the two group. After FDR correction, group differences were found in the time range from 338 to 356 ms (CI_{AVc} - AV_m vs. NH_{AVc} - AV_m), as indicated by the red and grey bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2.3. Erps to initial trials: distinct congruency-related ERP topographies in CI users and NH controls

ERPs to initial trials were compared between the two groups by means of the *global field power* (GFP) and the *global map dissimilarity* (GMD) to quantify group differences in ERP response strength and ERP topography, respectively (Murray et al., 2008). The GFP analysis (CI_{AVc} - AV_m vs. NH_{AVc} - AV_m) revealed no differences in ERP response strength between the CI users and the NH listeners (Fig. 4B). Regarding the GMD, the difference ERPs (CI_{AVc} - AV_m vs. NH_{AVc} - AV_m) showed distinct congruency-related ERP topographies between the two groups around 338 and 356 ms (Fig. 4C), which points to group-specific configurations of intracranial generators for the ERPs at this latency range (Lehmann and Skrandies, 1980; Murray et al., 2008). Specifically, the difference ERP of CI users (CI_{AVc} - AV_m) showed a frontal positivity around 338 and 356 ms – which corresponds to the offset part of the CRR response in CI users (Figs. 3, 4 and 5A). At the same latency range, however, the difference ERP of NH listeners (NH_{AVc} - AV_m) revealed a pronounced posterior positivity over posterior scalp regions, which corresponds to the onset part of the P3b response in NH listeners (NH_{AVc} - AV_m, Figs. 3, 4 and 5A).

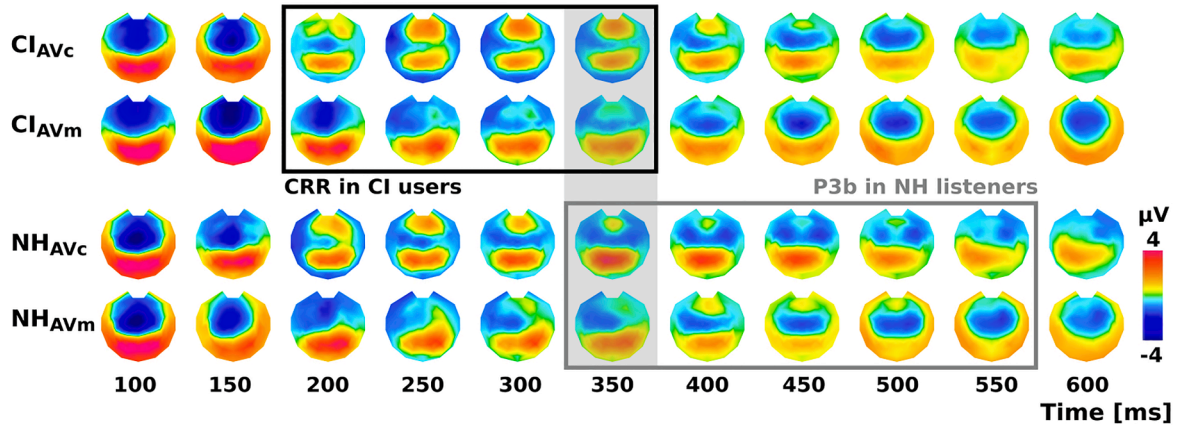
Taken together, these results indicate comparable ERP response strength (GFP) across groups but distinct congruency-related ERP topographies (GMD) from 338 to 356 ms between the CI users and the NH listeners. This suggests group-specific configurations of intracranial ERP generators during the perception of audio-visual objects at this latency range.

3.2.4. ERPs to initial trials: hierarchical clustering and single-subject fitting of differential ERP topographies

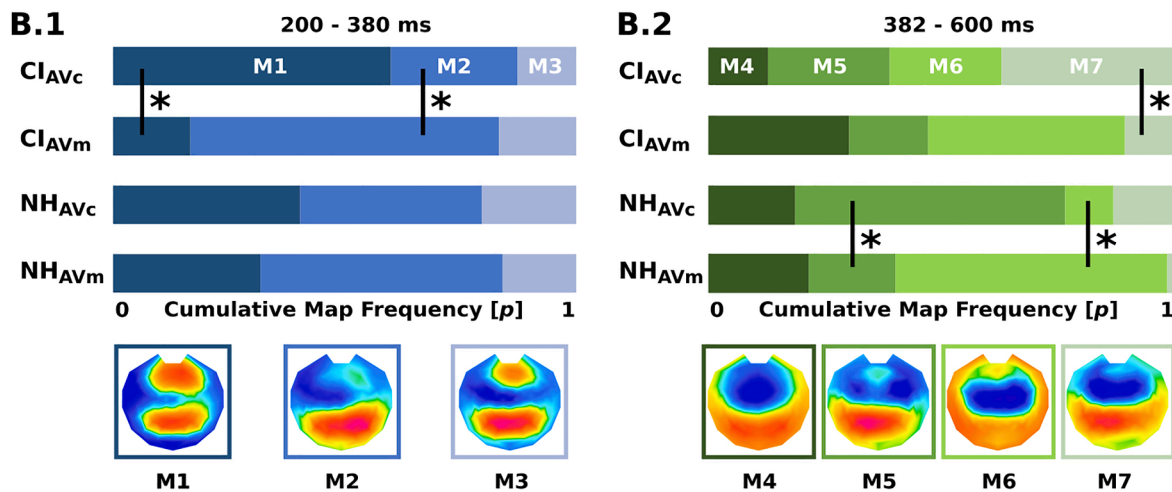
The hierarchical clustering of group-averaged data from AVc and AVm was conducted to ascertain both the temporal stability of ERP topographies across time as well as any differences across conditions and groups. This clustering identified 14 template maps in 16 clusters that collectively explained 90.3% of these concatenated data. Three circumscribed time windows were observed during the clustering of averaged post-onset ERP data (0–198, 200–380 and 382–600 ms). Identical template maps were observed during early post-sound onset periods until 198 ms between groups and conditions.

Over the 200–380 ms period, three template maps were identified (M1, M2, M3) that differentially characterized the ERPs between groups and conditions (Fig. 5B.1, see Supplementary Material C). Single-subject fitting was performed, and the total presence for each template map (M1-M3) was evaluated. The mixed-model repeated-measures ANOVA revealed a 3-way interaction ($F_{1.5,32.1} = 4.27, p = .033, \eta_p^2 = 0.16$), a Stimulus Condition \times Template Map interaction ($F_{1.5,32.1} = 12.2, p < .001, \eta_p^2 = 0.36$) and a main effect for Template Map ($F_{1.4,31.3} = 4.52, p$

A Voltage Topographies



B Topographic Clustering and Single-Subject Fitting



C Correlation Between d' and Single-Subject Fitting

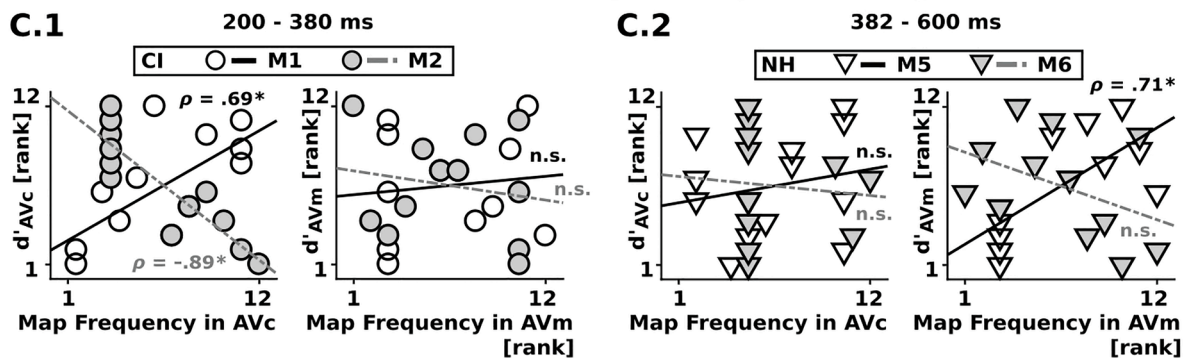


Fig. 5. Topographic clustering and single-subject fitting. A) Time course of voltage topographies for the initial audio-visual stimulus conditions AVc and AVm. Time windows of the CRR and the P3b ERPs are illustrated (black window: CRR, grey window: P3b; see also Fig. 4A.1 and 4A.3, respectively). The grey shade indicates topographies in the significant GMD time window (338–356 ms) of difference ERPs (CI_{AVc} - AVm vs. NH_{AVc} - AVm; see also Fig. 4C). B) Hierarchical clustering identified three template maps in the time window between 200 and 380 ms (B.1: M1, M2, M3) and four template maps in the time window between 382 and 600 ms (B.2: M4, M5, M6, M7). The number of samples that were assigned to each template map during single-subject fitting are depicted as cumulative map frequencies for maps M1-3 (blue, B.1) and M4-7 (green, B.2), for each group and condition separately. In CI users M1 and M2 dissociate between AVc and AVm stimulus conditions during the 200–380 ms time window (B.1). NH listeners show a dissociation of M5 (AVc) and M6 (AVm) during the 382–600 ms time window (B.2). C) Spearman rank correlations revealed for the CI users a positive relation between d'_{AVc} and the occurrence of M1 topographies for AVc, as well as a negative relation between d'_{AVc} and the occurrence of M2 topographies for AVm in the 200–380 ms time window (C.1). NH listeners showed a positive relationship between d'_{AVm} and the frequency of M5 for AVm in the 382–600 ms time window (C.2). Ranked values are depicted, corresponding to the Spearman correlations. Asterisks indicate significant statistical results ($p < .05$, Bonferroni-Holm corrected) in all panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

= .029, $\eta_p^2 = 0.17$). No other main effect or interaction was observed (all $p > .9$). In light of the significant 3-way interaction, separate ANOVAs were conducted for each group.

For CI users, a significant Stimulus Condition \times Template Map interaction was observed ($F_{1,4,15.2} = 17.5, p < .001, \eta_p^2 = 0.61$). No main effect was observed (all $p > .08$). Different patterns of template maps characterized ERPs in response to conditions AVc and AVm, respectively (Fig. 5B.1). Thus, in CI users the template map M1 better characterized responses to the AVc (number of samples 54.6 ± 10.3), compared to AVm (15.3 ± 5.4) stimulus condition ($t_{11} = 4.33, p = .002, r = 0.79$). At the same time, template map M2 better characterized responses to the AVm (60.5 ± 8.6) than AVc (24.8 ± 10.6) stimulus condition ($t_{11} = -4.65, p = .002, r = 0.81$). Template map M3 characterized both conditions equally ($p > .44$).

Neither main effect nor their interaction effect was reliable for NH listeners (all $p > .22$), suggesting a statistically indistinguishable pattern of template maps in response to both AVc and AVm stimulus conditions in NH listeners during the 200–380 ms time window.

Over the 382–600 ms period, four template maps were identified (M4, M5, M6, M7) that differentially characterized the ERPs between groups and conditions (Fig. 5B.2). Single-subject fitting was performed and the total number of map presence for each template map (M4–M7) was evaluated with a mixed-model repeated-measures ANOVA. A 3-way interaction ($F_{3,66} = 3.77, p = .015, \eta_p^2 = 0.15$) and a Stimulus Condition \times Template Map interaction ($F_{3,66} = 14.56, p < .001, \eta_p^2 = 0.4$) were revealed. No other main effect or interaction was observed (all $p > .19$). Again, in light of the significant 3-way interaction, separate ANOVAs were conducted for each group.

A significant Stimulus Condition \times Template Map interaction was observed for CI users ($F_{3,33} = 6.6, p = .001, \eta_p^2 = 0.38$). No main effect was observed (all $p > .79$). Only template map M7 showed a condition-specific difference, better characterizing the AVc stimulus condition (41.3 ± 11.3), compared to AVm ($12.1 \pm 5.4; t_{11} = 3.31, p = .028, r = 0.71$; Fig. 5B.2). No effect was observed for template maps M4, M5 or M6 (all $p > .08$). In NH listeners, a significant Stimulus Condition \times Template Map interaction was revealed ($F_{3,33} = 10.64, p < .001, \eta_p^2 = 0.49$). Whereas the template map M5 better characterized the AVc (63.2 ± 13.6) stimulus condition (AVm: $20.2 \pm 9.1; t_{11} = 3.46, p = .016, r = 0.72$), template map M6 better characterized the AVm (63.8 ± 13.6) stimulus condition (AVc: $11.4 \pm 7.1; t_{11} = -4.51, p = .004, r = 0.72$; Fig. 5B.2).

Spearman rank correlations were computed between discriminabilities (d'_{AVc}, d'_{AVm}) and the map frequencies for the template maps that dissociated between the two stimulus conditions AVc and AVm for CI users (M1, characterizing AVc; M2, characterizing AVm; Fig. 5C.1) and NH listeners (M5, characterizing AVc; M6, characterizing AVm; Fig. 5C.2). In CI users, a positive relationship between d'_{AVc} and the number of samples that were assigned to template map M1 ($\rho = 0.69, p = .04$), as well as a negative relationship to M2 ($\rho = -0.89, p < .001$) were revealed during the AVc stimulus condition (Fig. 5C.1). The same template maps were not related to d'_{AVm} in CI users (all $p > .9$). Note that the total number of samples corresponding to M1 is not independent of samples corresponding to M2. Thus, the relationship to one map might be directly related to an inverse relationship to the other map. However, M1 and M2 voltage topographies were related to d'_{AVc} , but not d'_{AVm} , indicating a significant behavior-ERP relationship only during the audio-visual congruent condition. Still, it is not possible to specifically conclude whether behavioral changes in d'_{AVc} were mainly related to the occurrence of M1 (positive correlation) or to M2 (negative correlation), or any combination of the two topographies. Please also note that reduced map counts in one condition might influence the validity of a correlation (see M2 during AVc in CI users, Fig. C.1, left panel). It is thus important to interpret the presented correlations only in synopsis, as it is done here. NH listeners showed a positive relationship between d'_{AVm}

and the template map M5 during the AVm stimulus condition ($\rho = 0.73, p = .038$; Fig. 5C.2), but no relationship was revealed between d'_{AVm} and M6 or between d'_{AVc} and neither template map (all $p > .5$).

These collective results reinforce the claim that CI users have a distinct processing of multisensory congruence. During the 200–380 ms time window, CI users exhibit distinct topographies (and by extension, distinct configurations of neural source) in response to congruent audio-visual stimuli, compared to audio-visual meaningless stimulus pairs (CI_{AVc}: M1, CI_{AVm}: M2; Fig. 5B.1). Critically, the presence of these voltage topographies during the AVc stimulus condition relates to the discriminability in the same condition (d'_{AVc} ; Fig. 5C.1). NH listeners do not exhibit distinct responses over this time window. However, during a subsequent time window (382–600 ms), topographic differences between stimulus conditions were observed in both CI users and NH listeners. Differences in response to congruent audio-visual stimuli dissociates in NH listeners (NH_{AVc}: M5, NH_{AVm}: M6; Fig. 5B.2), whereas the stimulus condition AVc is specifically characterized by template map M7 in CI users, but no dissociation with another template map was observed (Fig. 5B.2).

4. Discussion

We assessed the role of congruent audio-visual information on auditory recognition in elderly post-lingually deafened CI users and age-matched NH participants. Behavioral results (Fig. 2A) revealed a multisensory benefit for later auditory object recognition in both groups of subjects which was specific for the congruent audio-visual condition. The finding of such a congruency-specific multisensory facilitation effect suggests that the beneficial effect of multisensory stimulus context is driven by *congruent* visual information (benefit in stimulus condition AVc) rather than by the simple presence of additional visual information (no benefit in stimulus condition AVm). The EEG results (Figs. 3–5) revealed group differences in ERP topographies (at 338–356 ms) and a group-specific pattern of voltage topographies which were associated with (behavioral) auditory recognition ability between 200 and 380 ms in CI users and between 382 and 600 ms in NH listeners. In sum, our findings suggest that CI users have a distinct processing of multisensory congruence, which, however, allows both of these groups to remarkably improve the auditory recognition ability.

4.1. Audio-Visual context enhances auditory recognition

For young NH participants, Matusz and colleagues (2015) reported a significant improvement in auditory object recognition, if the respective objects had been associated previously with a congruent visual stimulus. This effect, in contrast, was not observed when the auditory objects were previously experienced in combination with a meaningless visual object or if they were perceived purely auditory. In the present study, the behavioral findings by Matusz and colleagues were extended by replicating the effect in both, an elderly sample of NH participants and a group of post-lingually deafened elderly CI users (Fig. 2). The observation of enhanced discriminability in the audio-visual congruent recognition condition suggests that both NH listeners and CI users integrated the respective visual and auditory naturalistic stimulus objects (Favre et al., 2014; Noel et al., 2015).

The recognition of auditory-only (d'_A) stimuli was impaired in CI users when compared to NH controls. Critically however, CI users reached the level of NH participants' auditory recognition when the initial auditory stimuli were paired specifically with congruent visual information (i.e., d'_{AVc} was comparable to NH d'_A , Fig. 2A). Importantly, the results of the sensitivity d' and the hit rates were consistent in that they revealed a beneficial impact on the auditory recognition ability, when these were preceded by congruent audio-visual objects (Fig. 2, Table 2). The results were further validated by the RTs of initial stimulus presentations: Both groups revealed faster RTs for audio-visual congruent information compared with audio-visual meaningless and

auditory-only stimuli (Fig. 2C.1). This is consistent with previous findings, showing that additional visual context improves auditory information processing with respect to accuracies (Matusz et al., 2015; Seitz et al., 2006), as well as RTs (Laurienti et al., 2004; Schierholz et al., 2017, 2015). Furthermore, previous EEG studies showed that congruent audio-visual perception modulates unimodal information processing already at the single-trial level (Matusz et al., 2015; Murray et al., 2005, 2004; Thelen et al., 2012; Thelen and Murray, 2013).

After implantation, CI users have to adapt to the electrical hearing, showing major cortical adaptations over the first few weeks and months after implantation (Green et al., 2008; Sandmann et al., 2015). Auditory rehabilitation typically approximates a plateau within the first year of CI experience (Holden et al., 2013; Lenarz et al., 2012). The present results indicate a positive effect of congruent audio-visual perception on the recognition of complex auditory objects, even in experienced CI users. Future studies will have to examine whether the reported effects also manifest in a long-term fashion. First results from NH participants with and without vocoded speech, as well as late-implanted deafened ferrets have already pointed to the potential effectiveness of audio-visual training protocols (Bernstein et al., 2014, 2013; Isaiah et al., 2014; Kawase et al., 2009; Seitz et al., 2006). Thus, we speculate that systematic audio-visual training protocols might yield an opportunity for CI users to improve recognition of auditory objects, even if this rehabilitation strategy may be applied several months after CI switch-on.

Our behavioral results were consistent in that they showed for both the CI users and the NH listeners a multisensory facilitation effect on the auditory recognition ability. In particular, the ANOVA computed on behavioral measures (hit rate, d') revealed no significant interaction effect between the factors Group (CI, NH) and Condition (AVm, A, AVc), which suggests a similar pattern of behavioral results across groups (see also section 3.1.1 for more details). However, at least on the descriptive level, CI users – when compared to NH listeners – revealed a relatively stronger improvement in the discriminability for the congruent audio-visual recognition condition (d'_A vs. d'_{AVc} and d'_{AVc} vs. d'_{AVm} ; Fig. 2). Indeed, supplementary analyses focusing on the *relative changes in performance* (A vs. AVc, A vs. AVm, AVc vs. AVm) confirmed that CI users – when compared to NH listeners – show a relatively stronger improvement in auditory object recognition ability when the initial auditory presentation was paired with semantically congruent visual information (Supplementary Material D). This finding is in line with previous literature, suggesting an enhanced interplay between the auditory and visual modality in CI users than in NH listeners (Giraud et al., 2001; Rouger et al., 2012, 2007; Schierholz et al., 2017, 2015; Strelnikov et al., 2015). The observed audio-visual enhancement in CI users can be explained by a compensatory process that CI patients might develop to overcome the limited CI input. Indeed, previous studies have reported that CI users show a strong bias towards the perception of the visual component of the multisensory stimulus, which points to altered processing and/or experience-related cortical changes in implanted individuals (Butera et al., 2018; Champoux et al., 2009; Rouger et al., 2007; Sandmann et al., 2012; Stevenson et al., 2017; Stropahl et al., 2017). However, based on the current behavioral results, we are reluctant to draw firm conclusions about the relative enhancement of the audio-visual gain in CI users. The supplementary results on the relative gains are based on a small sample size, and analyzing the relative gains bears the risk that the observed group effects are driven by baseline imbalances between the CI users and the NH listeners.

4.2. ERP activity indicates distinct audio-visual processing between groups

For the difference ERPs – reflecting the difference between the two audio-visual conditions (congruent vs. meaningless) – we observed a group effect in the ERP topographies for the time range between 338 and 356 ms ($CI_{AVc} - AVm$ vs. $NH_{AVc} - AVm$; Fig. 4). Specifically, the difference topographies of CI users showed an extended frontal Congruency-Related Response (CRR) when compared to NH participants (Figs. 3

and 4). By contrast, the topographies of NH participants revealed an earlier onset of the P3b component (Polich, 2007) as indicated by enhanced positivities over parieto-occipital scalp locations (Fig. 4). Furthermore, hierarchical clustering and single-subject fitting of template topographies revealed a topographical pattern in CI users during the 200–380 ms time window, dissociating the AVc and AVm stimulus conditions (Fig. 5B.1). At the same time, NH listeners showed dissociative topographies during the 382–600 ms time window (Fig. 5B.2). Given that different topographies indicate distinct configurations of neural sources (Lehmann and Skrandies, 1980; Murray et al., 2008), our results suggest that CI users show alterations in the processing of multisensory congruence when compared to NH listeners.

4.2.1. ERP topographies reflect perceptual mismatch and the reorientation of attention to visual information in CI users

The ERP results showed a pronounced frontal CRR response in CI users (Fig. 4A.1 and Fig. 5A), which highly corresponded to the M1 and M2 template maps that dissociated between the congruent and meaningless audio-visual conditions in the early time window (200–380 ms; Fig. 5B.1). Two different neuro-cognitive mechanisms may account for the more pronounced CRR response in CI users than in NH listeners (at the latency range 200–350 ms), in particular, 1) perceptual template matching/prediction and 2) the allocation of attentional resources.

Regarding the first mechanism, previous studies have reported on the anterior N2 ERP in the latency range between 200 and 350 ms (Folstein and Van Petten, 2007). Similarly, we observed in the present study a frontally distributed negativity – referred to as the CRR response – with more negative amplitudes during the meaningless audio-visual condition (template map M2) when compared to the congruent audio-visual condition (template map M1; Fig. 4A.1 and 5A). Furthermore, an early anterior ERP component in CI users, very similar to the CRR, was previously described in the context of an auditory oddball paradigm, indicating a mismatch of deviant syllables, compared to standard syllables (Soshi et al., 2014). The anterior N2 has been associated with general mechanisms of *matching sensory input with predictive template representations* (Folstein and Van Petten, 2007; Wang et al., 2001, 2004, 2003, 2002, 2000; Yang and Wang, 2002; Zhang et al., 2003). In this prediction framework, the N2 can be interpreted as signaling an error between the bottom-up sensation and the top-down expectation (van Veen et al., 2004). With regard to the present study, CI users may have experienced a conflict between the bottom-up sensation (of the limited CI input) and the top-down expectation (enabled by cortical representations), marked by slower responses (Nieuwenhuis et al., 2003; Zhang et al., 2003). This conflict may have been more pronounced in the audio-visual meaningless condition (AVm) than in the audio-visual congruent condition (AVc), because the random dot patterns in the meaningless condition did *not* provide supporting visual context. Thus, the more pronounced CRR response in CI users may be explained by the fact that in the audio-visual *meaningless* condition (AVm), the CI users had to rely on the ambiguous electrical input, which likely induced a mismatch or conflict between the auditory percept (bottom-up processing) and the experience-related cortical representations of auditory objects. According to Rönnberg and colleagues (2013), a perceptual mismatch changes perceptual processing from a fast, implicit information processing to a slower, explicit stimulus processing. In the present data, the slow responses in CI users (Fig. 2C) might reflect such an explicit stimulus processing, compared to NH participants, substantiating the notion of an increased perceptual mismatch in CI users during the ambiguous auditory recognition conditions (A, AVm). By contrast, in the *congruent* visual stimulus condition, CI users could reduce the ambiguity of the limited CI signal by considering the congruent visual information. This is indicated by enhanced behavioral performance (higher accuracies: d'_{AVc} , Fig. 2A.1), faster reaction times (AVc, Fig. 2C.1), and a *reduced* frontal ERP negativity ($CI_{AVc} - AVm$, Fig. 4A.1). Importantly, the single-subject fitting procedure revealed for the CI users an enhanced presence of the *M1 map* during the audio-visual *congruent* condition, and an

enhanced presence of the *M2 map* during the audio-visual *meaningless* condition at CRR latency (Fig. 5B.1). Thus, the anterior N2 might well correspond to our CRR response, which was observed by the topographical effects in the present data.

A second mechanism that may account for the more pronounced CRR response in CI users than in NH listeners is related to the *allocation of attentional resources*. Previous studies with NH listeners have suggested that the positive-going P2 (Crowley and Colrain, 2004) and the P3a ERP (with maximal positivity over frontal scalp regions; Polich, 2007) are related to *attention*. Specifically, the P3a may reflect the allocation of attention towards a behaviorally relevant object (Berti, 2016; Polich, 2007), whereas the P2 has been associated with the detection of informative cues, the disengagement of attention (García-Larrea et al., 1992), and the focused attention on a single modality (Mishra and Gazzaley, 2012). It is important to mention that the P2 and the P3a might overlap in time depending on the task (Berti, 2016). With regard to the present study, the pronounced CRR response in CI users might reflect the disengagement of attention from the auditory stream (P2) and/or the allocation of attention (P3a) towards the informative *congruent visual* information in the audio-visual congruent (AVc) stimulus condition (Butera et al., 2018). However, at this point we cannot exclude that the observed topographic differences in CI users are caused by altered visual-to-auditory priming (Schneider et al., 2008) or by cortical reorganization (Stropahl et al., 2017). Indeed, previous studies have shown altered processing of basic or more complex visual stimuli in CI users, which may be a consequence of sensory deprivation and/or degraded sensory input after implantation (Chen et al., 2016; Sandmann et al., 2012; Stropahl et al., 2015).

Taken together, the observed topographic differences (M1: AVc and M2: AVm, Fig. 5) and the extended CRR in CI users (Figs. 3 and 4A.1) may relate to the signaling of a prediction error in the audio-visual meaningless stimulus condition (anterior N2), and/or to the allocation of attention towards the visual modality in the audio-visual congruent stimulus condition (P2, P3a). Critically, the behavioral performance was specifically related to the template topographies only during the congruent audio-visual stimulus condition at the 200–380 ms time window (Fig. 5C.1; positive relation between d'_{AVc} and M1 and negative relation between d'_{AVc} and M2). This indicates that the CI users show a modulation of the CRR response in particular when an informative visual context is available.

4.2.2. P3b topographies reflect resource allocation in NH participants, but not in CI users

CI users and NH listeners showed distinct ERP difference topographies between 338 and 356 ms for audio-visual stimuli, dependent on the congruency of auditory and visual information ($CI_{AVc} - AVm$ vs. $NH_{AVc} - AVm$, Figs. 4 and 5). At this latency range, the NH listeners revealed specifically for the audio-visual *congruent* condition an ERP topography with a positivity over parieto-occipital scalp locations (Fig. 3F and Fig. 4A.3). Interestingly, this voltage topography is concordant with the template map M5 from the single-subject fitting procedure, which characterizes the audio-visual congruent condition specifically in the group of NH listeners in the later time window (382–600 ms). We suggest that this voltage topography of NH listeners in response to the audio-visual congruent condition (Fig. 4; M5 topography in Fig. 5B.2) reflects the P3b ERP (Polich, 2007). Together with the later offset of the CRR in CI users (Fig. 5A), a relatively early onset of the P3b in NH listeners (Fig. 5A) seems to drive the group-specific topographical differences ($CI_{AVc} - AVm$ vs. $NH_{AVc} - AVm$, Fig. 4).

Previous studies have associated the P3b with processing effort and the allocation of cortical resources in NH participants (Kok, 2001), as well as the accessibility of perceptual information in CI users (Henkin et al., 2009). Regarding the present study, the NH listeners showed a P3b response in the difference ERPs ($NH_{AVc} - AVm$; Fig. 3F and 4A.3), characterizing the audio-visual congruent condition during the 382–600 ms time window (template topography M5 in Fig. 5B.2). Here, the P3b

might reflect the reduced processing effort in NH listeners specifically when congruent visual context is available (Kok, 2001). This is confirmed by our behavioral results, showing increased discriminabilities (d'_{AVc}) and faster RTs (AVc, Fig. 2C.1) in particular in audio-visual congruent conditions.

In contrast to NH listeners, the CI users showed at the same latency range (382–600 ms) a template map M7 which characterized the audio-visual congruent condition. However, there was no dissociation to another template map in CI users (Fig. 5B.2). The reduced P3b in the difference ERP ($CI_{AVc} - AVm$, Fig. 4A.3) and the distinct template topographies during the AVc stimulus condition between NH listeners and CI users (NH: M5, CI: M7; Fig. 5B), point to a CI-induced resource capturing by the complex natural auditory stimuli (Finke et al., 2015). Although also CI users showed enhanced sensitivities (Fig. 2A) and faster RTs in the audio-visual congruent recognition and stimulus conditions (Fig. 2C.1), the reduced listening effort that is induced by congruent visual context might not lead to a reduced processing effort in total. Instead, CI users may allocate their resources to make increased use of the available congruent visual information, as indicated by alterations in congruent audio-visual processing at CRR latency. In contrast, NH listeners seem to use the same visual context information to complement and deepen their auditory perception (Rugg et al., 1998; Rugg and Yonelinas, 2003), as indicated by a pronounced P3b (Fig. 3F and 4) and behavioral improvements (Fig. 2).

5. Conclusion

A continuous recognition paradigm was employed to assess the facilitating potential of audio-visual perception on auditory object recognition in a group of elderly post-lingually deafened CI users and age-matched NH controls. An improved recognition ability of auditory stimuli was observed when these stimuli were preceded by congruent audio-visual context, both in the NH listeners and the CI users. Given that the observed performance gain in CI users might be of clinical relevance, future studies need to assess a) whether audio-visual strategies can improve the auditory rehabilitation after cochlear implantation and b) how audio-visual incongruent stimuli affect the observed processing differences in CI users.

Despite a similar pattern of behavioral improvements in both groups, our ERP results indicate a distinct congruency-related processing of audio-visual objects in CI users (before 350 ms) when compared to NH controls (after 350 ms). The distinctive ERP topographies in CI users may be explained by the processing of a perceptual mismatch (i.e., a conflict between the bottom-up and top-down processing) and/or by the enhanced allocation of attentional resources towards the available congruent visual information.

CRedit authorship contribution statement

Jan-Ole Radecke: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Visualization. **Irina Schierholz:** Investigation, Writing – original draft, Writing – review & editing. **Andrej Kral:** Writing – review & editing, Supervision, Resources. **Thomas Lenarz:** Writing – review & editing, Resources. **Micah M. Murray:** Methodology, Software, Visualization, Writing – review & editing. **Pascale Sandmann:** Conceptualization, Writing – original draft, Writing – review & editing, Resources.

Funding

This work was supported by the German Research Foundation (DFG) Cluster of Excellence “Hearing4all” [EXC 1077/1]. PS is supported by the DFG [SA 3615/1-1 and SA 3615/2-1] and by the Koeln Fortune Program/Faculty of Medicine, University of Cologne [278/2020]. MMM is supported by the Swiss National Science Foundation [169206].

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2022.102942>.

References

- Bar, M., 2007. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* 11 (7), 280–289. <https://doi.org/10.1016/j.tics.2007.05.005>.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7 (6), 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* <https://doi.org/10.2307/2346101>.
- Bernstein, L.E., Auer, E.T., Eberhardt, S.P., Jiang, J., 2013. Auditory perceptual learning for speech perception can be enhanced by audiovisual training. *Front. Neurosci.* 7, 1–16. <https://doi.org/10.3389/fnins.2013.00034>.
- Bernstein, L.E., Eberhardt, S.P., Auer, E.T., 2014. Audiovisual spoken word training can promote or impede auditory-only perceptual learning: results from prelingually deafened adults with late-acquired cochlear implants and normal-hearing adults. *Front. Psychol.* 5, 1–20. <https://doi.org/10.3389/fpsyg.2014.00934>.
- Berti, S., 2016. Switching attention within working memory is reflected in the P3a component of the human event-related brain potential. *Front. Hum. Neurosci.* 9, 1–10. <https://doi.org/10.3389/fnhum.2015.00701>.
- 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. <https://doi.org/10.1016/j.neuroimage.2014.02.031>.
- Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011, 1–15. <https://doi.org/10.1155/2011/813870>.
- Butera, I.M., Stevenson, R.A., Mangus, B.D., Woynaroski, T.G., Gifford, R.H., Wallace, M.T., 2018. Audiovisual temporal processing in postlingually deafened adults with cochlear implants. *Sci. Rep.* 8, 11345. <https://doi.org/10.1038/s41598-018-29598-x>.
- Champoux, F., Lepore, F., Gagné, J.-P., Théoret, H., 2009. Visual stimuli can impair auditory processing in cochlear implant users. *Neuropsychologia* 47 (1), 17–22. <https://doi.org/10.1016/j.neuropsychologia.2008.08.028>.
- 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.* 2016, 1–13. <https://doi.org/10.1155/2016/4382656>.
- 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. <https://doi.org/10.1016/j.clinph.2003.11.021>.
- Debener, S., Hine, J., Bleech, S., Eyles, J., 2008. Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology* 45, 20–24. <https://doi.org/10.1111/j.1469-8986.2007.00610.x>.
- 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.
- Drennan, W.R., Rubinstein, J.T., 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J. Rehabil. Res. Dev.* 45, 779–789. <https://doi.org/10.1682/JRRD.2007.08.0118>.
- Faivre, N., Mudrik, L., Schwartz, N., Koch, C., 2014. Multisensory integration in complete unawareness. *Psychol. Sci.* 25, 2006–2016. <https://doi.org/10.1177/0956797614547916>.
- Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., Sandmann, P., 2016. On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. *Neuropsychologia* 87, 169–181. <https://doi.org/10.1016/j.neuropsychologia.2016.05.019>.
- Finke, M., Sandmann, P., Kopp, B., Lenarz, T., Büchner, A., 2015. Auditory distraction transmitted by a cochlear implant alters allocation of attentional resources. *Front. Neurosci.* 9, 1–13. <https://doi.org/10.3389/fnins.2015.00068>.
- Folstein, J.R., Van Petten, C., 2007. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 6. <https://doi.org/10.1111/j.1469-8986.2007.00602.x>.
- Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.* 110 (2), 1150–1163. <https://doi.org/10.1121/1.1381538>.
- Friston, K., 2005. A theory of cortical responses. *Philos. Trans. R. Soc. London* 360 (1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>.
- García-Larrea, L., Lukaszewicz, A.-C., Manguière, F., 1992. Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia* 30 (8), 723–741. [https://doi.org/10.1016/0028-3932\(92\)90042-K](https://doi.org/10.1016/0028-3932(92)90042-K).
- Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* 16 (2), 129–135.
- Giraud, A.-L., Price, C.J., Graham, J.M., Truys, E., Frackowiak, R.S.J., 2001. Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron* 30 (3), 657–664. [https://doi.org/10.1016/S0896-6273\(01\)00318-X](https://doi.org/10.1016/S0896-6273(01)00318-X).
- Green, K.M.J., Ramsden, R.T., Julyan, P.J., Hastings, D.E.L., 2008. Cortical plasticity in the first year after cochlear implantation. *Cochlear Implants Int.* 9 (2), 103–117. <https://doi.org/10.1002/cii.358>.
- Hahlbrock, K.H., 1970. *Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet*. Thieme Verlag, Stuttgart.
- Härtig, C., Markowitsch, H.J., Neufeld, H., Calabrese, P., Deisinger, K., Kessler, J., 2000. *Wechsler Gedächtnis Test – Revidierte Fassung (WMS-R)*. Huber, Bern.
- Heimler, B., Weisz, N., Collignon, O., 2014. Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience* 283, 44–63. <https://doi.org/10.1016/j.neuroscience.2014.08.003>.
- Henkin, Y., Tetin-Schneider, S., Hildesheimer, M., Kishon-Rabin, L., 2009. Cortical neural activity underlying speech perception in postlingual adult cochlear implant recipients. *Audiol. Neurootol.* 14, 39–53. <https://doi.org/10.1159/000153434>.
- Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C., 2015. Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. *Audiol. Neurootol.* 19 (1), 21–26. <https://doi.org/10.1159/000371602>.
- Heydebrand, G., Hale, S., Potts, L., Gotter, B., Skinner, M., 2007. Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiol. Neurootol.* 12, 254–264. <https://doi.org/10.1159/000101473>.
- Holden, L.K., Finley, C.C., Firstz, J.B., Holden, T.A., Brenner, C., Potts, L.G., Gotter, B.D., Vanderhoof, S.S., Mispagel, K., Heydebrand, G., Skinner, M.W., 2013. Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* 34, 342–360. <https://doi.org/10.1097/AUD.0b013e3182741aa7>.
- Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
- House, W.F., Urban, J., 1973. Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Ann. Otol. Rhinol. Laryngol.* 82 (4), 504–517.
- Isaiah, A., Vongpaisal, T., King, A.J., Hartley, D.E.H., 2014. Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J. Neurosci.* 34 (33), 11119–11130. <https://doi.org/10.1523/JNEUROSCI.4767-13.2014>.
- Kawase, T., Sakamoto, S., Hori, Y., Maki, A., Suzuki, Y., Kobayashi, T., 2009. Bimodal audio-visual training enhances auditory adaptation process. *NeuroReport* 20, 1231–1234. <https://doi.org/10.1097/WNR.0b013e318232fbef8>.
- Klinke, R., Kral, A., Heid, S., Tillein, J., Hartmann, R., 1999. Recruitment of the auditory cortex in congenitally deaf cats by long-term cochlear electrostimulation. *Science* (80-) 285 (5434), 1729–1733.
- Kok, A., 2001. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38 (3), 557–577. <https://doi.org/10.1017/S0048577201990559>.
- Kollmeier, B., Wesselskamp, M., 1997. Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. Soc. Am.* 102 (4), 2412–2421. <https://doi.org/10.1121/1.419624>.
- Kral, A., 2007. Unimodal and cross-modal plasticity in the “deaf” auditory cortex. *Int. J. Audiol.* 46, 479–493. <https://doi.org/10.1080/14992020701383027>.
- Kral, A., Kronenberger, W.G., Pisoni, D.B., O’Donoghue, G.M., 2016. Neurocognitive factors in sensory restoration of early deafness: a connectome model. *Lancet Neurol.* 15 (6), 610–621. [https://doi.org/10.1016/S1474-4422\(16\)00034-X](https://doi.org/10.1016/S1474-4422(16)00034-X).
- Kral, A., Sharma, A., 2012. Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35 (2), 111–122. <https://doi.org/10.1016/j.tins.2011.09.004>.
- Kral, A., Tillein, J., Heid, S., Klinke, R., Hartmann, R., 2006. Cochlear implants: cortical plasticity in congenital deprivation. *Prog. Brain Res.* 283–402. [https://doi.org/10.1016/S0079-6123\(06\)57018-9](https://doi.org/10.1016/S0079-6123(06)57018-9).
- Kral, A., Yusuf, P.A., Land, R., 2017. Higher-order auditory areas in congenital deafness: top-down interactions and corticocortical decoupling. *Hear. Res.* 343, 50–63. <https://doi.org/10.1016/j.heares.2016.08.017>.
- 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, 405–414.
- Lee, T.-W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.* 11 (2), 417–441. <https://doi.org/10.1162/08997669900016719>.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48 (6), 609–621.
- Lehrl, S., 1999. *Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B*. Spitta.
- Lenarz, M., Sönmez, H., Joseph, G., Büchner, A., Lenarz, T., 2012. Cochlear implant performance in geriatric patients. *Laryngoscope* 122 (6), 1361–1365. <https://doi.org/10.1002/lary.23232>.
- 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. <https://doi.org/10.1038/nn.2653>.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., Moore, B.C.J., 2006. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *PNAS* 103 (49), 18866–18869. <https://doi.org/10.1073/pnas.0607364103>.
- Macmillan, N.A., Creelman, D.C., 2005. *Detection Theory: A User’s Guide*, 2nd ed. Lawrence Erlbaum Associates, Mahwah, NJ.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164 (1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.

- Matusz, P.J., Thelen, A., Amrein, S., Geiser, E., Anken, J., Murray, M.M., 2015. The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur. J. Neurosci.* 41 (5), 699–708. <https://doi.org/10.1111/ejn.12804>.
- Matusz, P.J., Wallace, M.T., Murray, M.M., 2017. A multisensory perspective on object memory. *Neuropsychologia* 105, 243–252. <https://doi.org/10.1016/j.neuropsychologia.2017.04.008>.
- Meyers, J.E., Meyers, K.R., 1995. *Rey Complex Figure Test and Recognition Trial-Professional Manual*. Psychological Assessment Resources.
- 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. <https://doi.org/10.1016/j.neuroimage.2017.11.062>.
- Mishra, J., Gazzaley, A., 2012. Attention distributed across sensory modalities enhances perceptual performance. *J. Neurosci.* 32 (35), 12294–12302. <https://doi.org/10.1523/JNEUROSCI.0867-12.2012>.
- Mosnier, I., Bebear, J.-P., Marx, M., Fraysse, B., Truy, E., Lina-Granade, G., Mondain, M., Sterkers-Artières, F., Bordure, P., Robier, A., Godey, B., Meyer, B., Frachet, B., Poncet-Wallet, C., Bouccara, D., Sterkers, O., 2015. Improvement of cognitive function after cochlear implantation in elderly patients. *JAMA Otolaryngol. Neck Surg.* 141, 1–9. <https://doi.org/10.1001/jamaoto.2015.129>.
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20 (4), 249–264. <https://doi.org/10.1007/s10548-008-0054-5>.
- Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. *J. Neurosci.* 26, 1293–1302. <https://doi.org/10.1523/JNEUROSCI.4511-05.2006>.
- Murray, M.M., De Lucia, M., Brunet, D., Michel, C.M., 2009. Principles of topographic analyses for electrical neuroimaging. In: Handy, T.C. (Ed.), *Brain Signal Analysis*. The MIT Press, Cambridge, MASS, USA, pp. 21–54.
- Murray, M.M., Foxe, J.J., Wylie, G.R., 2005. The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* 27 (2), 473–478. <https://doi.org/10.1016/j.neuroimage.2005.04.016>.
- Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., Schnider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21 (1), 125–135. <https://doi.org/10.1016/j.neuroimage.2003.09.035>.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118 (12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>.
- Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., Ridderinkhof, K.R., 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn. Affect. Behav. Neurosci.* 3 (1), 17–26. <https://doi.org/10.3758/CABN.3.1.17>.
- Noel, J.-P., Wallace, M., Blake, R., 2015. Cognitive neuroscience: integration of sight and sound outside of awareness? *Curr. Biol.* 25 (4), R157–R159. <https://doi.org/10.1016/j.cub.2015.01.007>.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118 (10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>.
- Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87.
- Rescorla, R.A., Solomon, R.L., 1967. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74 (3), 151–182.
- Rönner, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signorel, C., Stenfeld, S., Pichora-Fuller, M.K., Rudner, M., 2013. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7, 1–17. <https://doi.org/10.3389/fnsys.2013.00031>.
- 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. <https://doi.org/10.1002/hbm.21331>.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P., 2007. Evidence that cochlear-implanted deaf patients are better multisensory integrators. *PNAS* 104 (17), 7295–7300. <https://doi.org/10.1073/pnas.0609419104>.
- Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., Allan, K., 1998. Dissociation of the neural correlates of implicit and explicit memory. *Nature* 392 (6676), 595–598. <https://doi.org/10.1038/33396>.
- Rugg, M.D., Yonelinas, A.P., 2003. Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn. Sci.* 7 (7), 313–319. [https://doi.org/10.1016/S1364-6613\(03\)00131-1](https://doi.org/10.1016/S1364-6613(03)00131-1).
- 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, 555–568. <https://doi.org/10.1093/brain/awr329>.
- 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, 1967–1979. <https://doi.org/10.1093/brain/awp034>.
- Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Debener, S., Jäncke, L., Meyer, M., 2010. Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clin. Neurophysiol.* 121 (12), 2070–2082. <https://doi.org/10.1016/j.clinph.2010.04.032>.
- Sandmann, P., Plotz, K., Hauthal, N., de Vos, M., 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. <https://doi.org/10.1016/j.clinph.2014.06.029>.
- 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. <https://doi.org/10.1002/hbm.23515>.
- 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. <https://doi.org/10.1016/j.heares.2015.08.009>.
- Schneider, T.R., Debener, S., Oostenveld, R., Engel, A.K., 2008. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage* 42 (3), 1244–1254. <https://doi.org/10.1016/j.neuroimage.2008.05.033>.
- Schorr, E.A., Fox, N.A., van Wassenhove, V., Knudsen, E.I., 2005. Auditory-visual fusion in speech perception in children with cochlear implants. *Proc. Natl. Acad. Sci.* 102 (51), 18748–18750. <https://doi.org/10.1073/pnas.0508862102>.
- Seitz, A.R., Kim, R., Shams, L., 2006. Sound facilitates visual learning. *Curr. Biol.* 16 (14), 1422–1427. <https://doi.org/10.1016/j.cub.2006.05.048>.
- Sharma, A., Dorman, M.F., Spahr, A.J., 2002a. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539. <https://doi.org/10.1097/01.AUD.0000042223.62381.01>.
- Sharma, A., Dorman, M.F., Spahr, A.J., 2002b. Rapid development of cortical auditory evoked potentials after early cochlear implantation. *NeuroReport* 13 (10), 1365–1368. <https://doi.org/10.1097/00001756-200207190-00030>.
- Shinn-Cunningham, B.G., 2008. Object-based auditory and visual attention. *Trends Cogn. Sci.* 12 (5), 182–186. <https://doi.org/10.1016/j.tics.2008.02.003>.
- Simmons, F.B., Epley, J.M., Lummis, R.C., Guttman, N., Frishkopf, L.S., Harmon, L.D., Zwicker, E., 1965. Auditory nerve: electrical stimulation in man. *Science* (80-) 148 (3666), 104–106.
- Soshi, T., Hisanaga, S., Kodama, N., Kanekama, Y., Samejima, Y., Yumoto, E., Sekiyama, K., 2014. Event-related potentials for better speech perception in noise by cochlear implant users. *Hear. Res.* 316, 110–121. <https://doi.org/10.1016/j.heares.2014.08.001>.
- Stefanics, G., Kremláček, J., Czizler, I., 2014. Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8, 1–19. <https://doi.org/10.3389/fnhum.2014.00666>.
- Stevenson, R.A., Sheffield, S.W., Butera, I.M., Gifford, R.H., Wallace, M.T., 2017. Multisensory integration in cochlear implant recipients. *Ear Hear.* 38, 521–538. <https://doi.org/10.1097/AUD.0000000000000435>.
- Strelnikov, K., Rouger, J., Démonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2013. Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain* 136, 3682–3695. <https://doi.org/10.1093/brain/awt274>.
- Strelnikov, K., Rouger, J., Lagleyre, S., Fraysse, B., Démonet, J.-F., Déguine, O., Barone, P., 2015. Increased audiovisual integration in cochlear-implanted deaf patients: independent components analysis of longitudinal positron emission tomography data. *Eur. J. Neurosci.* 41 (5), 677–685. <https://doi.org/10.1111/ejn.12827>.
- Stropahl, M., Chen, L.-C., Debener, S., 2017. Cortical reorganization in postlingually deaf cochlear implant users: intra-modal and cross-modal considerations. *Hear. Res.* 343, 128–137. <https://doi.org/10.1016/j.heares.2016.07.005>.
- 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. <https://doi.org/10.1016/j.neuroimage.2015.07.062>.
- Thelen, A., Cappe, C., Murray, M.M., 2012. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage* 62 (3), 1478–1488. <https://doi.org/10.1016/j.neuroimage.2012.05.027>.
- Thelen, A., Matusz, P.J., Murray, M.M., 2014. Multisensory context portends object memory. *Curr. Biol.* 24 (16), R734–R735. <https://doi.org/10.1016/j.cub.2014.06.040>.
- Thelen, A., Murray, M.M., 2013. The efficacy of single-trial multisensory memories. *Multisens. Res.* 26 (5), 483–502. <https://doi.org/10.1163/22134808-00002426>.
- Thelen, A., Talsma, D., Murray, M.M., 2015. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition* 138, 148–160. <https://doi.org/10.1016/j.cognition.2015.02.003>.
- van Veen, V., Holroyd, C.B., Cohen, J.D., Stenger, V.A., Carter, C.S., 2004. Errors without conflict: implications for performance monitoring theories of anterior cingulate cortex. *Brain Cogn.* 56 (2), 267–276. <https://doi.org/10.1016/j.bandc.2004.06.007>.
- van Wassenhove, V., Grant, K.W., Poeppel, D., 2007. Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45 (3), 598–607. <https://doi.org/10.1016/j.neuropsychologia.2006.01.001>.
- Viola, F.C., Debener, S., Thorne, J., Schneider, T.R., 2010. Using ICA for the analysis of multi-channel EEG data. In: Ullsperger, M., Debener, S. (Eds.), *Simultaneous EEG and fMRI: Recording, Analysis and Application*. University Press, Oxford, pp. 121–133.
- Viola, F.C., Thorne, J.D., Bleeck, S., Eyles, J., Debener, S., 2011. Uncovering auditory evoked potentials from cochlear implant users with independent component analysis. *Psychophysiology* 48, 1470–1480. <https://doi.org/10.1111/j.1469-8986.2011.01224.x>.
- Wang, H., Wang, Y., Kong, J., Cui, L., Tian, S., 2001. Enhancement of conflict processing activity in human brain under task relevant context. *Neurosci. Lett.* 298 (3), 155–158. [https://doi.org/10.1016/S0304-3940\(00\)01757-2](https://doi.org/10.1016/S0304-3940(00)01757-2).
- Wang, Y., Cui, L., Wang, H., Tian, S., Zhang, X., 2004. The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology* 41, 21–29. <https://doi.org/10.1111/j.1469-8986.2003.00134.x>.

- Wang, Y., Kong, J., Tang, X., Zhuang, D., Li, S., 2000. Event-related potential N270 is elicited by mental conflict processing in human brain. *Neurosci. Lett.* 293 (1), 17–20. [https://doi.org/10.1016/S0304-3940\(00\)01480-4](https://doi.org/10.1016/S0304-3940(00)01480-4).
- Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., Zhang, X.i., 2003. Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Exp. Brain Res.* 148 (4), 451–457. <https://doi.org/10.1007/s00221-002-1319-y>.
- Wang, Y., Wang, H., Cui, L., Tian, S., Zhang, Y., 2002. The N270 component of the event-related potential reflects supramodal conflict processing in humans. *Neurosci. Lett.* 332 (1), 25–28. [https://doi.org/10.1016/S0304-3940\(02\)00906-0](https://doi.org/10.1016/S0304-3940(02)00906-0).
- Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data – a practical approach. *J. Neurosci. Methods* 250, 34–46. <https://doi.org/10.1016/j.jneumeth.2014.08.002>.
- Wilson, B.S., Dorman, M.F., 2008. Cochlear implants: a remarkable past and a brilliant future. *Hear. Res.* 242 (1-2), 3–21. <https://doi.org/10.1016/j.heares.2008.06.005>.
- Yang, J., Wang, Y., 2002. Event-related potentials elicited by stimulus spatial discrepancy in humans. *Neurosci. Lett.* 326 (2), 73–76. [https://doi.org/10.1016/S0304-3940\(02\)00204-5](https://doi.org/10.1016/S0304-3940(02)00204-5).
- Zhang, X.i., Wang, Y., Li, S., Wang, L., 2003. Event-related potential N270, a negative component to identification of conflicting information following memory retrieval. *Clin. Neurophysiol.* 114 (12), 2461–2468. [https://doi.org/10.1016/S1388-2457\(03\)00251-7](https://doi.org/10.1016/S1388-2457(03)00251-7).