Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing

Maren Stropahl a,⁎, Karsten Plotz b, Rüdiger Schönfeld b, Thomas Lenartz c,d, Pascale Sandmann d,e, Galit Yovel f, Maarten De Vos d,g,h, Stefan Debener a,d

⁎ Corresponding author at: Department of Psychology, European Medical School, Carl von Ossietzky University Oldenburg, Germany.

E-mail address: maren.stropahl@uni-oldenburg.de (M. Stropahl).

Article history:
Received 19 February 2015
Accepted 21 July 2015
Available online 26 July 2015

Keywords:
Cochlear implant
Source analysis
Visual face processing
Lip reading
Cross-modal reorganization

ABSTRACT

There is converging evidence that the auditory cortex takes over visual functions during a period of auditory deprivation. A residual pattern of cross-modal take-over may prevent the auditory cortex to adapt to restored sensory input as delivered by a cochlear implant (CI) and limit speech intelligibility with a CI. The aim of the present study was to investigate whether visual face processing in CI users activates auditory cortex and whether this has adaptive or maladaptive consequences. High-density electroencephalogram data were recorded from CI users (n = 21) and age-matched normal hearing controls (n = 21) performing a face versus house discrimination task. Lip reading and face recognition abilities were measured as well as speech intelligibility. Evaluation of event-related potential (ERP) topographies revealed significantly higher activation in the right auditory cortex for CI users compared to NH controls, confirming visual take-over. Lip reading skills were significantly enhanced in the CI group and appeared to be particularly better after a longer duration of deafness, while face recognition was not significantly different between groups. However, auditory cortex activation in CI users was positively related to face recognition abilities. Our results confirm a cross-modal reorganization for ecologically valid visual stimuli in CI users. Furthermore, they suggest that residual takeover, which can persist even after adaptation to a CI is not necessarily maladaptive.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

It is known that sensory-deprived brain regions do not remain inactive but that missing unimodal sensory input results in cortical changes (Merabet and Pascual-Leone, 2009). Recent work on visual and auditory deprivation showed converging evidence of cross-modal reorganization after a time of deprivation. In the case of human deafness, the auditory cortex seems to take over visual functions (Finney et al., 2003; Finney et al., 2001; Karns et al., 2012) and this take-over has been related to enhanced visual abilities (Bavelier et al., 2000; Bavelier et al., 2006; Hauthal et al., 2013; Lomber et al., 2010). Cross-modal plasticity can have adaptive and maladaptive effects (Heimler et al., 2014) and may therefore influence the degree of auditory rehabilitation with a cochlear implant (CI).

Previous studies have shown that not only the developing brain (Sharma et al., 2005; Sharma et al., 2007) but also the mature brain of middle-aged (35–62 years) and elderly CI recipients (74–78 years) rapidly adapts to the partly restored (electrical) input within the first weeks after initial implant use (Sandmann et al., 2014). This adaptation process may partly indicate a reversal of deafness-induced loss of functional specialization (Giraud et al., 2001; Pantev et al., 2006; Sandmann et al., 2014; Sharma et al., 2007), and partly reflect the adaptation to the coarse, artificial input as provided by a CI. However, the performance level in hearing and speech comprehension varies strongly among CI users. This suggests differences in the capacity of the auditory cortex to adapt to the electrical input signal after implantation. Pre- and post-surgical factors are known to influence the individual benefit of the CI, among them the onset of hearing loss, the duration of deafness, the extent of residual hearing and CI experience (Blamey et al., 1996,
The experience of auditory deprivation is thought to induce a visual take-over type of reorganization in the auditory cortex which is not completely reversed after implantation. Insufficient adaptation to the new input may be reflected by residual signs of visual take-over (Doucet et al., 2006; Lee et al., 2001; Sandmann et al., 2012). Accordingly, it was found that a residual cross-modal take-over is maladaptive which is reflected in an inverse relation to the speech recognition ability with a CI (Buckley and Tobey, 2011; Doucet et al., 2006; Sandmann et al., 2012). Several studies with deaf individuals and CI users have shown that the effect of deprivation-induced cross-modal plasticity has mostly been localized to the right hemisphere (Cardin et al., 2013; Doucet et al., 2006; Finney et al., 2001; Rouger et al., 2012; Sandmann et al., 2012), either because the right hemisphere is more susceptible to reorganization than the left hemisphere (Lazard et al., 2013) or because the right hemisphere is more involved in the processing of sounds with low complexity (Hine and Deebener, 2007). How cortical reorganization affects visual abilities of CI users is not yet thoroughly investigated. The aim of this study was therefore to further investigate cross-modal reorganization in CI users and its consequences, by using ecologically valid visual stimuli.

Human face perception has been studied intensively in the past years. Faces, as compared to inanimate objects, are perceived in a specialized manner (Kanwisher, 2000). Face-selectivity has been verified in neuro-imaging studies which have identified the occipital face area (OFA) and the fusiform face area (FFA) as core regions in a neural network of face processing (Haxby et al., 2000; Kanwisher et al., 1997; Kanwisher and Yovel, 2006). Face-selectivity can also be observed in electrophysiological responses, in particular the N170 component (approx. 170 ms after face onset) which typically shows the largest amplitudes over occipito-temporal scalp regions in the right hemisphere (Rentin et al., 1996; Bötzel and Grüsser, 1989; Rossion and Jacques, 2008). The N170 component is larger for faces compared to other objects like houses (Rossion and Jacques, 2008). A well-known effect in the domain of face processing is the face-inversion-effect (Eimer, 2000; Haxby et al., 1999; Kanwisher et al., 1998; Sadeh and Yovel, 2010; Valentine, 1988). Inverted (rotated 180° to upside-down) faces are processed differently than upright faces, as revealed by larger N170 amplitudes and delayed latencies (~8 ms; Eimer, 2000) to inverted faces. By contrast, the inversion effect is much less pronounced for other objects (Rossion et al., 2000; McPartland et al., 2004) observed a trend in their study that a faster neural processing speed (assessed by the N170 latency) reflects a better face recognition ability in healthy participants. Nevertheless, not much is known about the relation between the neurophysiology of visual face processing and behavioral correlates. To the best of our knowledge no studies investigating this relation have been conducted with CI users so far.

There is evidence that deaf individuals have advantages in face processing. This is reflected in a more accurate matching of faces (Arnold and Murray, 1998; De Heering et al., 2012) and may have evolved because the deaf focus more intensively on faces in order to compensate the missing auditory input during face-to-face social communication (Kral et al., 2013; Mitchell et al., 2013; Woodhouse et al., 2009). At present it is not well understood whether this advantage applies also to CI users (Rouger et al., 2012). Previous studies have suggested superior lip reading abilities and different patterns of visual language processing in CI users when compared to normal-hearing controls (Giraud and Truy, 2002; Lee et al., 2007; Rouger et al., 2007, 2012). We therefore hypothesized that CI users show superior abilities in face recognition (assessed by the Cambridge Face Memory Test) and lip reading. It is expected that lip reading skills are increased with a longer duration of deafness due to broader experience in lip reading. On the neurophysiological level, we expected advantages in visual face processing abilities in CI users, as is reflected in enhanced neural activity (larger N170 component) or faster processing speed (shorter N170 latency). Furthermore, by mean of distributed source modeling we determined whether CI users show the predicted activation of the auditory cortex during the processing of visually presented face. We investigated whether this cross-modal reorganization has maladaptive consequences on the individual CI benefit, which should be reflected in lower speech perception.

2. Materials and methods

2.1. Participants

Twenty-one post-lingually deafened individuals implanted with a cochlear implant (13 women, 8 men) participated in the study. The participants showed a variety of hearing-loss etiologies and were all unilaterally implanted at the time of testing (Table 1). Fourteen of the CI users got the implant on the right side and seven on the left side. All CI users had been using their implant for at least 12 months on a regular basis, which is approximately 16 hours a day. The CI experience varied between 12 and 187 months (M = 54.8, SE = 9.1 months). The duration of severe hearing loss/deafness was subjectively reported by the participants. We defined the duration of deafness based on the time at which the participants could not benefit from hearing aids anymore which was mirrored in very insufficient speech recognition, until the date of implant surgery. The duration of deafness ranged from three to 240 months (M = 88.7, SE = 19.3 months). The age at onset of hearing loss varied between a very early onset at birth and an acquired hearing loss during adulthood (range: 0–51 years, mean and standard error: M = 16.4, SE = 3.6 years). Even if the onset of hearing loss was very early in life for some of the participants, the actual onset of profound deafness was always after speech acquisition, which is also reflected in relatively high speech intelligibility scores. None of the CI users had active sign language skills. Additionally, a normal-hearing (NH) control group, matched with the CI users in gender and age, was tested. The mean age of the CI group was M = 51.1, SE = 3.6 years (range: 20–74 years) and M = 50.1, SE = 3.6 of the NH group (range: 21–74 years). For further analyses, the hearing score in noise and in silence was measured from each participant at the date of investigation. A standard pure-tone audiogram with headphones was measured to ensure the normal-hearing status of the control group. The boundary of the age-appropriate hearing level was set to ~30 dB HL for the frequencies 500 Hz, 1 kHz, 2 kHz, 3 kHz and 4 kHz. The study was conducted in accordance with the local ethical committee guidelines of the University of Oldenburg and in agreement with the declaration of Helsinki. Every participant gave written informed consent before the onset of the experiment.

2.2. Face vs. house discrimination task

The participants performed a face vs. house discrimination task. Images of faces and houses were shown in randomized order on a computer screen, spanning a visual field of 6° horizontally and vertically (Fig. 1). The face pictures were taken from the Harvard Face Database and pictures of houses were taken from the website http://www.zoo.map.co.uk/. All pictures were equally scrambled, gray scaled and optically matched in contrast and luminance. The undegraded stimulus material was successfully used in a previous study investigating N170 face processing (De Vos et al., 2012). The pictures could either be upright or inverted (rotation of 180°) which yielded four different picture categories. Before each trial, a fixation cross of 600 ms was presented. After a uniform jitter lasting between 0 and 1000 ms, a picture was displayed for 250 ms. Pictures were followed by a grey screen presented for approximately 1700 to 2500 ms, indicating a valid response interval. Trial duration was 3 seconds in total. The whole experiment consisted of three blocks with a total duration of 22 minutes. Twelve different pictures were used in each category. A block consisted of 108 pictures in
total (27 per category), presented in random order. There was a break of 60 seconds in between the blocks. The participants were instructed to respond as fast and as accurate as possible with the corresponding button press to either a face (upright or inverted) or a house (upright or inverted). The button for faces and houses was equally randomized to either the right or the left thumb over all participants. Each participant accomplished a practice trial before starting the recording. Response times as well as accuracies for each participant and each stimulus category were computed.

2.3. EEG recording

EEG data were collected with a BrainAmp EEG amplifier system (BrainProducts, Gilching, Germany). A cap with a 96 Ag/AgCl electrode layout, equidistantly placed with a central frontopolar site as ground and a nose-tip reference was used (Easycap, Herrsching, Germany). This infra-cerebral spatial sampling covers more of the head sphere than traditional 10–20 electrode layouts, and supports source localization efforts (e.g. Hauthal et al., 2014; Hine and Debener, 2007; Hine et al., 2008; Sandmann et al., 2014). Two electrodes were placed below the eyes to control for eye blinks and eye movements. Data was recorded with a sampling rate of 500 Hz applying an online analog filter from 0.1 to 250 Hz. Electrode impedances were kept below 20 kΩ. The stimulus presentation was controlled with the Presentation software (Neurobehavioral Systems, Albany, CA, USA). The CI users took their speech processor off during data acquisition to preserve the data from any electromagnetic distortion.

2.4. EEG data analysis

EEG data were preprocessed with EEGLAB 11.0.5.4b (Delorme and Makeig, 2004) in the MATLAB environment (Mathworks). After initial 1 Hz high pass filtering and a joint probability-based artifact correction (3 SD) for training the algorithm, an independent component analysis (ICA) based on the extended Infomax (Bell and Sejnowski, 1995; Jung et al., 2000a; Jung et al., 2000b) was applied. The resulting unmixing weights were used to linearly decompose the original raw data and attenuate typical artifacts such as eye blinks, eye movements and
electrical heart-beat artifacts (Viola et al., 2009). The ICA-corrected data was filtered between 0.1 and 30 Hz and segmented into epochs from −200 to 800 ms. Epochs were corrected to a 200 ms pre-stimulus baseline. Remaining artifactual epochs were identified and rejected using a joint probability measure (standard deviation: 4). The resulting data quality for each single participant was visually confirmed. The mean number of trials available for averaging was, for all conditions, M = 255.5 (SE = 21) for the CI users and M = 254 (SE = 20.2) for the NH controls (t < 1, n.s.). The analysis focused on the face-selective N170 component of the event-related potential (ERP). Usually, ERP studies investigate the effect of the face-selective N170 component focus on electrodes over lateral occipito-temporal sites, such as P08 and PO10 in the international 10–20 system (Rossion and Jacques, 2008). We therefore selected the three electrodes of our electrode layout in closest approximation to P08, PO8 and PO10. Only electrodes located over the right hemisphere were selected, since rightward lateralization for face processing has been previously reported (Campanella et al., 2000; Kanwisher and Yovel, 2006; Rossion et al., 2003). Furthermore, based on earlier results we hypothesized that the CI users show a visual take-over of the auditory cortex which usually is particularly pronounced in the right hemisphere (Buckley and Tobey, 2011; Finney et al., 2003; Lazard et al., 2013; Sandmann et al., 2012). Therefore, to reduce the number of statistical tests, we restricted the subsequent analyses to the right hemisphere.

For N170 quantification, the individual N170 amplitude was defined as the average of a 4 ms window around the most negative deflection within the time interval of 140 to 200 ms at electrodes P08, PO8 and PO10 (Sadeh et al., 2008). The latency of the N170 component was derived as the time of the most negative peak in the same time window. Analyses of variance (ANOVA) for repeated measures were used to analyze N170 amplitudes and latencies. The 2 × 2 × 3 × 2 way ANOVA comprised the factors object (face, house), object orientation (upright, inverted) and electrode (P8, PO8 and PO10) as within-subject factors, and group (CI, NH) as between-subjects factor. The significance threshold was set to p < .05 (two-tailed). Effect sizes are indicated with $\eta^2_p$ (partial eta squared). For further correlation analyses, a face selectivity index as proposed previously (Sadeh et al., 2010) was computed for each participant. This index normalizes the difference between the electrophysiological strength of the N170 component for houses and faces by the peak-to-peak amplitude of the P100 and the N170 face component to reveal a quantification of face-selectivity strength $(N170_{face} − N170_{house}) / (P100_{face} − N170_{house})$. The face index reflects the overall amplitude of the signal to faces. Due to the naturally negative deflection of the N170 ERP, the index becomes more negative with stronger responses to faces.

In a next step, we extended the analysis beyond the classical grand-average ERP ANOVA model. Specifically we investigated spatio-temporal differences between conditions and groups since the predicted visual-take over mechanism may alter N170 voltage maps in CI users. To test this assumption, individual general linear models (GLM) as implemented in the linear modeling (LIMO) toolbox were used (Pernet et al., 2011). The GLM was computed for each subject and included the data from all 96 electrodes. Previous studies have reported differences between CI users and NH controls not only over occipital regions but also over more anterior scalp regions, typically located over the right hemisphere (Cardin et al., 2013; Doucet et al., 2006; Finney et al., 2003; Lazard et al., 2013; Sandmann et al., 2012). The GLM approach allowed us to test for the hypothesis of a group difference in more anterior scalp regions. Based on the estimated GLM parameters, a two-sample t-test between the CI group and the NH group was computed for each of the 96 electrodes at each time point for the time interval of the N170 component (140–200 ms), separately for the different conditions (upright faces, inverted faces, upright houses, inverted houses). A temporal cluster correction procedure was applied to account for multiple comparison errors resulting from a high number of statistical comparisons. For the nonparametric statistical test a map of clusters is calculated using permutation, given a pre-defined threshold (p < 0.05). Cluster mass is computed based on the sum of the statistical values of the cluster. Statistical tests and cluster maps are as well computed for the observed data which are then compared to the thresholds (cluster sums) obtained from the permutation test (Maris and Oostenveld, 2007). Corrected p-values are reported.

To address the issue of implantation side in the group of CI users (14 right and 7 left), a $2 × 2 × 3 × 2$ way ANOVA was computed for the CI group comprising the factors object (face, house), object orientation (upright, inverted) and electrode (P8, PO8 and PO10) as within-subject factors, and with the between-subject factor CI-side for the electrophysiological data.

Behavioral data, such as response times and accuracies were analyzed with a $2 × 2 × 2$ repeated-measures ANOVA, including object (face, house) and object orientation (upright, inverted) as within-subject factors and group (CI, NH) as between-subjects factor.

2.5. Source analysis

The freely available Brainstorm software was used to estimate the contribution of visual and auditory cortex to the face-selective N170 component (Tadel et al., 2011). Specifically, we applied the method of dynamic statistical parametric mapping (dSPM, Dale et al., 2000) to identify and evaluate active sources. The dSPM method calculates the estimated locations of the scalp-recorded electrical activity of the neurons by minimum-norm inverse maps, with constrained dipole orientations. The activity is normalized by an individual estimate of the noise standard deviation at each location (Hämäläinen and Ilmoniemi, 1984). Similar to minimum norm solutions, dSPM exactly localizes sources, but with blurred spatial resolution (Lin et al., 2006). The boundary element method (BEM) implemented in OpenMEEG was used as a forward solution. It provides three realistic layers and therefore representative anatomical information (Gramfort et al., 2010; Stenroos et al., 2014). In a first step, dSPM source estimation was calculated for each of the four conditions separately based on 96-channel ERP data referenced to a common average. Secondly, based on our hypotheses, regions of interest (ROI) for visual and auditory areas were defined prior to statistical computations. Source activation in these ROIs were statistically compared between CI users and NH controls. The ROIs were defined based on the Destrieux-atlas implemented in Brainstorm (Destrieux et al., 2010; Tadel et al., 2011), which relies on automatic parcellation using a surface-based alignment of the cortical folding (Destrieux et al., 2010). The visual ROI comprised partly the fusiform face area, which is known to be actively involved in face processing (Kanwisher and Yovel, 2006). The visual ROI also encompassed the medial occipito-temporal and the lingual sulci (Destrieux: S_oc-temp_med_and_lingual). Regarding the auditory cortex, previous studies have reported cross-modal activity in Brodmann areas 41 and 42 of deaf individuals and CI users (Bottari et al., 2014; Cardin et al., 2013; Finney et al., 2003; Sandmann et al., 2012). Similarly, we defined the auditory ROI as a combination of three small regions to get a close approximation to Brodmann areas 41 and 42 (Destrieux: G_temp_sup-G_T_transv, S_temporal_transverse and G_temp_sup-Plan_tempo). To address concerns about spatially unspecific differences between CI users and NH controls, a control ROI was defined at the occipital pole, closely related to BA 17, the primary visual cortex (Destrieux: pole_occipital). Consistent with the ERP analysis, the source analysis focused on the time window of the N170 component (140–200 ms). The individual absolute peak magnitude of the ROI source activation was defined as the average of a 4 ms window around the peak and was subjected to statistical analyses. Where appropriate, two-sample t-tests between the groups were applied for each time
point within the N170 time interval. P-values were corrected for multiple comparison errors.

2.6. The Cambridge Face Memory Test

To assess face recognition abilities, the participants completed the upright version of the Cambridge Face Memory Test (Bowles et al., 2009; Duchaine and Nakayama, 2006) in a separate session. The Cambridge Face Memory Test (CFMT) is a quick clinical administration for prosopagnosia being comprised of facial learning and short-term recognition of faces. The test consists of 72 items in total and is separated in three subsets of different difficulty levels. The stimuli only show facial aspects of the individuals and no other components like hair or clothes to avoid general picture information. Only moderate correlations with other visual memory tasks (abstract art $r = .26$, Wilmer et al., 2010; or cars $r = .37$, Dennett et al., 2012) indicate its sensitivity for face memory. The CFMT has a high internal reliability (Cronbach’s alpha $= .89$, Bowles et al., 2009). The normal population covers a wide range of scores without showing a ceiling effect, which is optimal for the use of the CFMT in potentially high performance group. The clinical cut-off for prosopagnosia are scores two standard deviations lower than the mean from the norm population (Bowles et al., 2009; Duchaine and Nakayama, 2006). The CFMT is known to show the “other-face-effect” by declined means for memory and perceptual discrimination if the observed faces are not of the participant’s race (Bowles et al., 2009). According to the literature, the mean score of the CFMT for German participants is 72.2% (52 out of 72 items) within the age range of 18 to 35 years (Bowles et al., 2009; Herzmann et al., 2008). Bowles et al. (2009) also showed that age might influence the results of the CFMT. The mean recedes for a normal population older than 50 years. Unfortunately, no norm data are available for an elderly German population. To the best of our knowledge, no data is available for CI users that accomplished the CFMT. Due to withdrawal from further participation, the result of one CI user could not be evaluated. The correct recognition rate (accuracies in percent correct) of the upright version of the CFMT was assessed, and a one-tailed two-sample t-test was conducted to reveal the predicted face recognition advantage of CI users.

2.7. Lip reading quantification

Participants performed an additional behavioral lip reading task. The stimuli were monosyllabic words taken from a German speech recognition test (“Freiburger Sprachtest”; Hahlbrock, 1970), which is widely used for diagnostic purposes. The words were recorded in our lab as audio-visual videos with a professional speaker. For the present task only the visual component was used. The onsets of the spoken word was fixed to one second after video onset. The video resolution was 1920 × 1080 Pixel (Full HD). The test consisted of 31 words. Ten of these words were used at the beginning of the test for a feedback-based training. The test list consisted of 21 items. To quantify lip reading abilities, participants performed two experimental blocks in which the same test list was presented with different word order. The participants were seated in front of a screen watching the video without the audio signal. The task was to enter the perceived word by a keyboard. Due to technical malfunctioning the data from one NH control participant were lost. For the analysis, the percent of correctly reported words was assessed. Apart from obvious mistyping mistakes (e.g. interchanged letters) only completely correct words were considered in this score. Lip reading scores were subjected to a one-tailed t-test between groups to assess the prediction that CI users show superior lip reading performance when compared to age-matched NH listeners.

Significant ANOVA effects were followed up with two-tailed t-tests, unless directional predictions were tested. One-tailed tests are indicated where appropriate. Possible associations between behavioral and electrophysiological results were investigated with Pearson’s product-moment correlation analyses.

3. Results

3.1. Behavioral data

Due to the simplicity of the discrimination task, a ceiling effect was reached ($M = 97\%$, $SE = 0.41$). The $2 \times 2 \times 2$ ANOVA with accuracy as dependent variable revealed neither a main effect of group nor significant interactions with group ($F < 1$, n.s.). The mean response time across all subjects and conditions was $486$ ms ($SE = 15$). The $2 \times 2$ repeated-measures ANOVA revealed a significant main effect of object, $F_{1,39} = 16.66$, $p < .001$, $\eta^2 = .30$, and a main effect of object orientation, $F_{1,39} = 37.77$, $p < .001$, $\eta^2 = .49$, as well as a significant interaction between the two within-subject factors, $F_{1,39} = 7.79$, $p = .008$, $\eta^2 = .17$. Participants revealed shorter response times for upright faces ($M = 468$ ms, $SE = 15$) compared to upright houses ($M = 494$ ms, $SE = 15$). Additionally, delayed response times for inverted faces compared to upright faces were found (mean difference upright and inverted faces: $M = 14$ ms, $SE = 2$). The difference in response times to upright and inverted houses was not significant (mean difference upright and inverted houses: $M = 4$ ms, $SE = 2$). Furthermore, a trend for a main effect of group was found, $F_{1,39} = 3.0$, $p = .091$, $\eta^2 = .07$, indicating that CI users responded slower than NH controls (CI: $M = 512$ ms ($SE = 20$); NH: $M = 461$ ms ($SE = 95$)). To investigate individual differences and the effect of duration of deafness on response times, a Pearson’s product-moment correlation analysis revealed that a longer duration of deafness was related to longer response times to upright faces ($r_{18} = .447$, $p = .048$). This effect is not confound with the age of the participants, as the partial correlation that controlled for the potentially confounding age variable still revealed significance, $r_{19} = .521$, $p = .022$. The correlation between the response times for upright houses and duration of deafness was not significant.

The analysis of face recognition showed scores in a normal range (CFMT $M = 72.23\%$, $SE = 2.1$) and no significant differences between the CI users and the NH listeners ($t_{19} = 1.29$, $p = .10$, one-tailed). In the lip reading task, CI users showed a score of $38\%$ ($SE = 3.81$) and better lip reading than NH controls, who achieved a score of $27\%$ ($SE = 4.3$), $t_{19} = 1.99$, $p = .027$ (one-tailed). In order to assess the better lip reading abilities, correlation analyses were computed. A significant correlation between better lip reading skills and a longer duration of deafness was observed, $r_{19} = .569$, $p = .007$, as well as with the age at hearing loss onset, $r_{19} = -.586$, $p = .005$.

3.2. Event-related potentials (ERPs)

Fig. 2 shows the grand average ERP for the NH group (2A) and the CI group (2B) for the four conditions (faces, inverted faces, houses, inverted houses) for an electrode in close proximity to P010. As expected, the ERPs revealed for all conditions two distinct peaks, one positive peak around 100 ms, representing early visual processing (P100), and a negative deflection at approximately 170 ms, referred to as N170.

The $2 \times 2 \times 2 \times 2$ repeated measures ANOVA with N170 amplitude as dependent variable yielded a significant object main effect, $F_{1,40} = 63.52$, $p < .001$, $\eta^2 = .61$. As expected, participants elicited larger N170 amplitudes to faces ($M = −8.44 \mu V$, $SE = 1.17$) compared to houses ($M = −3.43 \mu V$, $SE = 0.94$). The main effect of object orientation was highly significant as well, $F_{1,40} = 10.89$, $p = .002$, $\eta^2 = .21$, and was driven by larger N170 amplitudes for upright compared to inverted objects (mean difference upright and inverted faces: $M = 1.78 \mu V$, $SE = 0.39$; mean difference upright and inverted houses: $M = 0.19 \mu V$, $SE = 0.18$). For this analysis, no significant main or interaction effects emerged with group ($F < 1$, n.s.). A significant three-way interaction of object, object orientation and electrode was found, $F_{1,40} = 11.65$, $p < .001$, $\eta^2 = .23$, and followed up with separate $2 \times 2$ ANOVAs for each electrode. For the follow-up ANOVA the significant main effect of object, indicating larger N170 amplitudes for faces compared to houses,
was confirmed at all three electrode sites (P8: \(F = 57.65, p < .001, \eta^2_p = .58\); P08: \(F = 46.65, p < .001, \eta^2_p = .53\); PO10: \(F = 75.4, p < .001, \eta^2_p = .65\)). The main effect of object orientation was significant as well for electrode P08, \(F = 8.83, p = .005, \eta^2_p = .18\), and PO10, \(F = 24.91, p < .001, \eta^2_p = .38\). Electrode PO10 additionally revealed a significant interaction of object and object orientation, \(F = 44.1, p = .001, \eta^2_p = .26\). A post-hoc t-test showed larger N170 amplitudes for upright faces compared to inverted faces, \(t_{41} = −4.78, p < .001\), with a mean difference of 1.9 \(\mu V (SE = 0.4)\). No inversion effect was observed for houses.

It may be argued that the implantation side could have an influence on the reported results. A separate 2x2x3x2 ANOVA for the CI users was applied to investigate differences based on implantation side. The electrophysiological data between CI users that have the CI on the left or on the right side did not show any significant difference. However, this null finding is based on small and unequal sample sizes. Furthermore, at present, the literature in the field does not allow to isolate laterality from effects based on the implanted ear.

### 3.3. Spatio-temporal analysis

The extended analysis using the GLM approach addressed the hypothesis of a topographic group difference. The two-sample t-test by time point t-test revealed six electrodes located over right occipito-temporal scalp regions which showed significantly larger negative deflections to upright faces for the CI users compared to NH listeners in the time of the N170 component (with \(p < .048\) as determined by the temporal cluster correction). Among these significant electrodes, the position of one electrode was located at a lateral occipito-temporal region close to PO10, whereas the position of the remaining electrodes was located more anterior over right temporal areas. Fig. 3 shows the t-color-coded values for the significant electrodes and time points, which spanned a latency range comprising the N170. Additionally, the scalp distribution of t-values is shown at the latency of the largest group difference (166 ms). The group comparison revealed a more anterior, right lateralized distribution showing stronger responses to faces for the CI users compared to the NH controls. None of the remaining scalp regions, latencies or experimental conditions (inverted faces, upright and inverted houses) revealed a significant group difference.

### 3.4. Face-selectivity analysis

Analysis of the N170 face-selectivity index, which quantifies differences between N170 amplitude for upright faces and houses, did not reveal any significant difference between CI users and NH controls. To test for a higher face-selectivity based on a longer duration of auditory deprivation, a correlation analysis between the duration of deafness and face selectivity was computed. The analysis confirmed that CI users with a longer duration of deafness revealed a stronger face-selectivity compared to participants with a shorter duration of deafness (Fig. 4B; electrode PO10: \(r_{19} = −.398 p = .037\), one-tailed). Furthermore, for both groups we tested whether face recognition was related to the strength of the face-selective N170 component. Only the NH group showed a positive relationship between face recognition and face-selectivity (electrode PO8: \(r_{18} = −.472 p = .016\) (one-tailed); electrode PO10: \(r_{18} = −.443 p = .022\), one-tailed).

### 3.5. N170 latency analysis

A 2×2×3×2 repeated-measures ANOVA for N170 latency as dependent variable revealed a significant object main effect, \(F_{1,40} = 27.83, p < .001, \eta^2_p = .41\). Showing that houses (\(M = 160.86\) ms, \(SE = 2.51\)) had a lower latency than faces (\(M = 171.84\) ms, \(SE = 2.37\)). The main effect of object orientation was significant as well, \(F_{1,40} = 16.79, p < .001\).
p < .001, $\eta^2 = .30$, showing delayed latencies for inverted ($M = 167.58$ ms, SE = 2.22) compared to upright objects ($M = 165.13$ ms, SE = 2.25). To investigate the inversion effect, post-hoc t-tests locating the significant interaction effect of object and object orientation, $F_{1,40} = 10.76, p = .002, \eta^2 = .21$, were computed. The tests revealed shorter latencies for upright compared to inverted faces with a mean difference of $M = −.44$ ms, SE = .90 ($t_{41} = −4.87, p < .001$), whereas the difference between upright and inverted houses was not significant ($t < 1$, mean difference $M = .5, SE = .77$). The main factor electrode was also significant, $F_{1,40} = 10.83, p < .001, \eta^2 = .21$. Subsequent t-tests revealed that peak latencies at electrode P8 were on average 3.6 ms delayed compared to electrodes PO8 and PO10 ($P8$ vs. PO8 $t_{41} = 4.01, p < .001$; P8 vs. PO10 $t_{41} = 3.43, p = .001$).

3.6. Source analysis

Peak magnitudes of evoked cortical responses to visual face stimuli were compared between the groups in a time interval of 140 to 200 ms within the two distinct ROIs, encompassing face-selective (visual ROI) and auditory processing areas (auditory ROI). The latency at the peak of activity in the respective ROI was collected for each participant. For both groups, time series are shown for the visual (Fig. 5) and the auditory (Fig. 6) ROI for upright faces and houses. The overall time series of activation pattern showed for both groups a peak of activation around 170 ms in the visual ROI (Fig. 5C). A similar pattern was seen for the auditory ROI (Fig. 6C). This was evident only for the face but not for the house conditions. Moreover, CI users showed descriptively a larger peak compared to the NH controls in the visual as well as in the auditory ROI. Based on the finding that the face-inversion effect did not differ between groups and was furthermore not at all seen for houses, the analysis of the source activity focused only on the conditions of upright faces and houses.

Regarding activation of the visual ROI, a 2 × 2 repeated-measures ANOVA with magnitude as dependent variable and the factors object (upright faces, houses) and group was computed. A main effect object identified a highly significant effect, $F_{1,40} = 171.2, p < .001, \eta^2 = .81$, showing higher activity for faces compared to houses (faces: $M = 96, SE = .05$; houses: $M = 39, SE = .04$). The same ANOVA model was applied to magnitude peak latencies and showed a highly significant object main effect, $F_{1,40} = 6.17, p = .017, \eta^2 = .13$, indicating shorter latencies for houses than faces (faces: $M = 173.24, SE = 2.8$; houses: $M = 166.52, SE = 3.18$). The interaction between object and group, $F_{1,40} = 3.17, p = .083, \eta^2 = .07$, showed a trend for differences in latencies to faces and houses between CI users and NH controls. A post-hoc analysis revealed that latency of houses and faces differed significantly for the NH controls ($t_{20} = 2.87, p = .01$) with shorter latencies for houses ($M = 163.43, SE = 4.4$) compared to faces ($M = 174.95, SE = 4.81$). The CI users did not show a significant difference in the latencies for face and house conditions. No group difference was observed.

Analysis of the auditory ROI magnitude was done with a 2 × 2 ANOVA model using object as within-subject factor (houses, faces) and group (CI, NH) as between-subject factor. The main factor object was highly significant, $F_{1,40} = 74.83, p < .001, \eta^2 = .65$, and the interaction between object and group was marginally significant, $F_{1,40} = 4.02, p = .052, \eta^2 = .09$. Following up this interaction (Fig. 7) revealed higher auditory cortex activity for the CI users compared to the NH controls, but this was evident only for faces ($t_{20} = −2.011, p = .025$, one-tailed) and not for houses ($t_{20} = −.04, p = .49$). To identify the time course of this effect, one-tailed t-tests between the groups were applied for each time point within the N170 time interval (140 within the N170 time interval 200 ms). CI users showed significantly higher activity compared to the NH controls in the auditory ROI between 150 and 194 ms (see Fig. 6, grey shaded area). Interestingly, the CI users also showed a trend for a positive relationship between activity in the auditory ROI and the face recognition ability ($r_{18} = .413 p = .071$), which was absent in the NH controls ($r_{18} = .193 p = .401$). The activation pattern elicited by houses did not differ between groups (Fig. 7). The magnitude of the auditory ROI and the lip reading ability did not reveal a significant
correlation for the CI users and for the NH controls (CI: \( r_{19} = -0.109 \) \( p > 0.1 \), NH: \( r_{19} = -0.084 \) \( p > 0.1 \)).

The 2 × 2 ANOVA for the peak activity latencies did not reveal any significant effect. Numerically, the CI users revealed shorter latencies in the auditory ROI to faces compared to the NH controls (CI: \( M = 171.43 \) ms, SE: 2.6; NH: \( M = 178.1 \) ms, SE: 4.3). Peak latency was correlated to behavioral performance, as revealed by Pearson’s product-moment correlations with the CFMT as well as lip reading results. The shorter latencies for the CI users in the auditory ROI were related to a higher accuracy in lip reading (Fig. 4D: \( r_{19} = -0.468 \) \( p = 0.032 \)), suggesting a functional correlate of auditory cortex activation by visually presented faces. This correlation was not significant for the NH controls. Correlations of the peak latency with the CFMT results were not significant for either the CI user or the NH controls (\( p > 0.1 \)). To address the issue of a generally higher global activity for CI users, an independent two-sample t-test for the upright face condition was computed between the groups in the control ROI. These analyses did not show a significant group difference (magnitude: \( t_{38} = 0.56 \) \( p > 0.1 \); latency: \( t_{38} = 0.45 \) \( p > 0.1 \)) within the time interval of the N170. Descriptively, the CI users showed less activity at a latency of 100 ms.

4. Discussion

This study investigated face processing in post-lingually deafened cochlear-implant (CI) users. Unfortunately CI users cannot undergo functional magnetic resonance imaging. Alternative technologies such as...
as functional near infrared spectroscopy may identify visual take-over as well but have not yet been fully exploited in the context of CI research (Chen et al., 2015; Lawler et al., 2015). Therefore, we used risk-free high-density EEG recordings combined with distributed source localization to identify cortical activity in CI users and carefully matched NH individuals. Compared to NH controls, CI users showed significantly larger negative deflections to faces over more anterior right temporal scalp regions, and subsequent source analysis confirmed the predicted pattern of enhanced auditory cortex activation in this group. These results add to the growing body of literature reporting a visual take-over of the auditory cortex in hearing-impaired individuals. In CI users, visual activation in the auditory cortex cannot only be observed for low level visual stimuli but also for more complex, ecologically relevant face stimuli. Moreover, behavioral analysis of lip reading and face recognition abilities in CI users suggested a functional role of auditory cortex activity, since better face recognition was associated with larger auditory cortex activation and better lip reading was associated to shorter auditory cortical latencies at the time of the N170 component. As expected, lip reading skills were better in CI users than in NH controls and appeared to be particularly enhanced after an earlier onset of hearing loss and a longer duration of deafness.

4.1. Intra-modal reorganization

On the scalp level, CI users had larger responses to faces over more anterior regions, and on the source level they showed higher activation of occipito-temporal cortical sources, albeit the latter effect was not statistically significant. A significant correlation for CI users to be more face-selective after a longer period of auditory deprivation was found as well. This suggests that the difference in strength of their electrophysiological response to faces and houses becomes more pronounced with a longer duration of deafness. These findings support the view that changes due to auditory deprivation do not only occur in auditory areas but can affect also the visual cortex of CI users (Buckley and Tobey, 2011; Champoux et al., 2009; Doucet et al., 2006; Turgeon et al., 2012). This is in line with the idea that enhanced brain responses in visual areas reflect an intra-modal reorganization. Evidence in favor of this idea has been reported for CI users (Buckley and Tobey, 2011; Doucet et al., 2006; Giraud et al., 2001a; Rouger et al., 2007), congenitally deaf individuals (Armstrong et al., 2002; Neville and Lawson, 1987) and even individuals with mild to moderate hearing impairment (Campbell and Sharma, 2014).

Contradictory to this pattern of enlarged VEPs, Sandmann et al. (2012) recently reported smaller VEP amplitudes for CI users. Several reasons may account for this discrepancy. First, differences in results might be related to different characteristics of the stimuli used, that is, low-level revers ing checkerboard patterns of different luminance ratios in the previous study (Sandmann et al., 2012) and more complex, ecologically valid static faces in the current study. Furthermore, in the current study we are looking at the face-specific N170 component, while Sandmann et al. (2012) focused on the P100 component which is highly influenced by physical stimulus properties. A second reason may be that the task was fundamentally different between the two studies. In the checkerboard study (Sandmann et al., 2012), the participants fixated the center of the checkerboard images, by contrast, in the current study, the house vs. face discrimination task required an active behavioral response. Thus, discrepancies in the results on CI users may be at least partially related to the processing resources being differentially allocated at the sensory and cognitive level for the different tasks as suggested by Gardin et al. (2013). Finally, the two groups of CI users studied differed in their duration of deafness, the duration of CI experience and, importantly, in their implants. Sandmann et al. tested a mixed group of uni- and bilaterally implanted participants, whereas in our study we restricted our analyses to unilaterally implanted CI users. At present it is not clear if and how cortical changes are influenced by a second CI in late deaf individuals.

4.2. Functional changes due to intra-modal reorganization

Even though intra-modal reorganization has been associated with superior visual functionality in deaf people (Bottari et al., 2010; Neville and Lawson, 1987), in the present study a task-related functional superiority for CI users, such as faster response times to face stimuli, could not be found. To the contrary, there was a trend of slower response times for the CI users which were associated to longer a duration of deafness. To the best of our knowledge, studies testing visual response times for post-lingually deafened adult CI users are rare. Gilley and colleagues used a visual detection task and found that CI users showed slower response times when compared to normal hearing controls (Gilley et al., 2010). Interestingly, this group difference was observed only for late implanted children, while early implanted children were not slower than NH children, confirming the view of improved cortical function among children receiving implants at early ages (Eggermont and Ponton, 2003; Kirk et al., 2002; Tyler et al., 2000). Regarding deaf individuals, previous results suggest that changes in visual abilities are stimulus- and task-selective, and thus may be revealed only under specific conditions (Pavani and Bottari, 2012). Whereas deaf individuals outperformed the NH controls in a direction of motion discrimination task, an overall group difference could not be found for a movement localization task (Hauthal et al., 2013). Furthermore, a study by De Heering et al. (2012) showed slower response times for deaf individuals in matching upright faces compared to normal hearing controls. In contrast, Bottari et al. (2010) found faster response times for deaf individuals in a visual detection task. Thus, the superior functionality of deaf individuals cannot be generalized and seems selective for attentionally demanding visual tasks (Bavelier et al., 2006; Pavani and Bottari, 2012). This interpretation might hold for CI users as well. Further studies investigating non-linguistic face processing are needed to find out whether an intra-modal reorganization in CI users might have functional consequences concerning face detection or even face recognition.

4.3. Cross-modal reorganization

The processing of ecologically relevant static face stimuli requires higher order visual functions and, as predicted, elicited a pattern of cross-modal reorganization in CI users. This complements previous findings on CI users reporting visual activation of auditory cortex during the processing of simple visual stimuli (Sandmann et al., 2012). Accordingly, the recruitment of auditory areas could be found in studies
showing enlarged as well as reduced VEPs in the visual cortex (Buckley and Tobey, 2011; Campbell and Sharma, 2014; Doucet et al., 2006; Giraud, Price, et al., 2001; Sandmann et al., 2012), which suggests a more complex interaction between intra-modal (visual) and cross-modal (auditory) reorganization resulting from auditory deprivation. We addressed the concern that the higher activity in the visual as well as the auditory ROI for the CI users may originate from a stronger global activation by investigating a control ROI (related to the primary visual area), which is clearly involved in visual processing. The group difference was not significant within the time interval of the N170. Descriptively, a trend indicated smaller amplitudes in the control ROI for CI users compared to NH controls at 100 ms latency. This analysis further supports our main finding of a regionally and functionally specific cross-modal reorganization in CI users. What exactly visual activation of auditory areas in CI users indicates, remains to be understood. Previous studies suggested an enhancement of the dorsal visual stream by cross-modal reorganization, as deaf individuals showed functionally better motion processing and an enlarged peripheral vision compared to NH individuals (Armstrong et al., 2002; Bavelier et al., 2000; Lomber et al., 2010; Neville and Lawson, 1987). Nevertheless, it is still debated whether dorsal or ventral visual stream functions are more prone to cross-modal reorganization (Armstrong et al., 2002; Campbell and Sharma, 2014; Lomber et al., 2010; Vachon et al., 2013). Face processing can be regarded as a specialty of ventral stream processing. Hence, our results of a combination of higher activation in visual face processing as well as auditory areas in CI users compared to NH controls might imply neural network modifications of ventral stream processing (Weisberg et al., 2012). An enhanced ventral pathway processing might be a compensatory mechanism caused by a stronger need to focus on facial information to understand speech and conversation-related information (Campbell and Sharma, 2014; Kral et al., 2013; Streltsov et al., 2013). Although previous findings seem to be ambiguous, arguments for an enhancement of either the ventral or the dorsal pathway might not be mutually exclusive. It is likely that the dichotomic vision of reorganization in either the ventral or the dorsal pathway is oversimplified and that cross-modal plasticity possibly relates to an interaction between sensory and functional reorganization, affecting both visual pathways (Armstrong et al., 2002; Campbell and Sharma, 2014; Cardin et al., 2013; Vachon et al., 2013).

4.5. Influence of pre- and post-surgical factors

Studies are often conducted with relatively small groups of CI users, and participants are heterogeneous concerning pre- and post-surgical factors due to difficulties in the recruitment of participants (Campbell and Sharma, 2014; Doucet et al., 2006; Sandmann et al., 2012). Important pre-operative factors that could influence the benefit of a CI are e.g. the age at onset of severe hearing loss/deafness, the etiology of the hearing loss and also the residual hearing of both ears. The CI benefit could also be influenced post-operatively for example by the duration of CI experience (Blamey et al., 2012; Lazard et al., 2012). While our sample size of 21 participants is larger compared to previous studies investigating cross-modal reorganization in CI users, it is composed of a high heterogeneity concerning pre- and post-operative factors, which is also reflected in a relatively large variability of hearing performance. In previous studies a link between the strength of the visual-take over and CI benefit has been demonstrated (Buckley and Tobey, 2011; Doucet et al., 2006; Sandmann et al., 2012). Therefore, further studies with large and homogeneous groups of CI users are needed to investigate functional differences due to cross-modal reorganization. Ideally this sample should also be characterized by the state of adaptation of the auditory cortex to electrical input, as reflected in late auditory evoked potentials (Debener et al., 2008; Viola et al., 2012; Viola et al., 2011). How intra-modal (auditory) input relates to cross-modal (visual) take-over in the auditory cortex remains to be understood.

4.6. Behavioral measurements of face recognition and lip reading

We assessed face recognition using a popular test, the CFMT. The CFMT is known to be influenced by several factors (Bowles et al., 2009; Duchaine and Nakayama, 2006). The ethnicity effect known as “other-race-effect” is well established for face memory (Meissner and Brigham, 2001) and affects the results for e.g. German participants by reducing their scores compared to race-matched stimuli. The validated mean score for German participants testing their face recognition ability has been found to be 72.2 % which is approximately 8% reduced compared to participants from countries were the stimulus-ethnicity match is better (Bowles et al., 2009). The overall results of our sample (72.23 %) are very well in line with the validated norm scores of the CFMT. Since our groups were carefully matched in gender and age, these factors could not have biased group results for the CFMT. In contrast to our predictions, an overall better face recognition ability could not be found for the CI group. Unfortunately, to the best of our knowledge, studies investigating non-linguistic face processing in CI users do not exist. Our predictions were based on indirect evidence, since superior face recognition abilities have been found in deaf individuals. However, even in deaf individuals, the superiority for face processing compared to NH individuals is not well understood (De Heering et al., 2012). Most studies have shown that differences in face recognition are based on sign language use and not on non-linguistic face processing. The difference is mostly between deaf individuals and hearing non-signers, whereas hearing signers also perform better than hearing non-signers (Arnold and Murray, 1998; Bettag et al., 1997; De Heering et al., 2012; McCullough and Emmorey, 1997).

McPartland et al. (2004) found a trend for healthy individuals that faster neural processing is related to better face recognition. This finding could not be replicated in the present study. However, a correlation analysis using the face-selectivity index (Sadéh et al., 2008), which expresses relation between the amplitude for faces and non-face objects, identified a clear relationship to face recognition abilities in the predicted direction, namely, higher face-selectivity was associated with better face recognition. Interestingly, this relationship was observed for NH controls but not for CI users. Since the higher N170 amplitudes to faces in CI users did not reflect better face recognition abilities, they might index other visual processing stages not studied in the present experiment, or simply reflect alterations in visual network properties.
unrelated to task processing performance. As a consequence, CI users might not generally be better in face processing but may be more specialized in ecologically important abilities such as lip reading, which still includes face perception. This interpretation is supported by the results of our second behavioral test in which the CI users revealed significantly higher accuracies in lip reading compared to NH controls. The lip reading ability, as our results suggest, depends on the duration of deafness and, consequently, on the individual experience in lip reading. Our findings are consistent with other studies, showing that CI users used visual speech cues more efficiently than NH listeners and that they maintain this superior ability after longer time of CI experience (Rouger et al., 2007, 2012; Strelnikov et al., 2010).

5. Conclusion

The present EEG study investigated face processing of CI users in comparison to NH controls. Enhanced N170 ERPs and a higher cortical activation of occipito-temporal face processing areas for CI users indicated a pattern of intra-modal reorganization. Furthermore, a higher activation of auditory processing areas could be found in CI users reflecting a pattern of cross-modal reorganization elicited by complex and ecologically important face stimuli. The greater activation and the shorter latencies in the auditory cortex were positively related to the face recognition and the lip reading ability of the CI users. In general, our results suggest a pattern of functionally specialized cortical reorganization in CI users for ecologically valid visual stimuli after a time of auditory deprivation and sensory recovery, which is not necessarily maladaptive.

Acknowledgements

This research has received funding by the grant I-175-105.3-2012 of the German-Israel Foundation and by the cluster of excellence Haering4all Oldenburg, Germany.

References


Bavelier, D., Neville, H.J., 2002. Cross-modal reorganization in cochlear implant users for ecologically valid visual stimuli after a time of auditory deprivation and sensory recovery, which is not necessarily maladaptive.


