Brain Systems Mediating Voice Identity Processing in Blind Humans

Cordula Hölig,^{1,2*} Julia Föcker,³ Anna Best,¹
Brigitte Röder,¹ and Christian Büchel²

¹*Department of Biological Psychology and Neuropsychology, University of Hamburg, Germany*

²*Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany*

³*Department of Psychology and Educational Sciences, University of Geneva, Switzerland*

Abstract: Blind people rely more on vocal cues when they recognize a person's identity than sighted people. Indeed, a number of studies have reported better voice recognition skills in blind than in sighted adults. The present functional magnetic resonance imaging study investigated changes in the functional organization of neural systems involved in voice identity processing following congenital blindness. A group of congenitally blind individuals and matched sighted control participants were tested in a priming paradigm, in which two voice stimuli (S1, S2) were subsequently presented. The prime (S1) and the target (S2) were either from the same speaker (person-congruent voices) or from two different speakers (person-incongruent voices). Participants had to classify the S2 as either a old or a young person. Person-incongruent voices (S2) compared with person-congruent voices elicited an increased activation in the right anterior fusiform gyrus in congenitally blind individuals but not in matched sighted control participants. In contrast, only matched sighted controls showed a higher activation in response to person-incongruent compared with person-congruent voices (S2) in the right posterior superior temporal sulcus. These results provide evidence for crossmodal plastic changes of the person identification system in the brain after visual deprivation. *Hum Brain Mapp* 35:4607–4619, 2014. © 2014 Wiley Periodicals, Inc.

Key words: congenitally blind; sensory deprivation; plasticity; voice; person recognition; identity; functional magnetic resonance imaging

INTRODUCTION

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: Federal Ministry of Education and Research; Contract grant number: G01GW0561.

*Correspondence to: Cordula Hölig, Biological Psychology and Neuropsychology, University of Hamburg, Von-Melle-Park 11, 20146 Hamburg, Germany. E-mail: cordula.hoelig@uni-hamburg.de
Received for publication 12 July 2013; Revised 10 February 2014; Accepted 13 February 2014.

DOI 10.1002/hbm.22498

Published online 17 March 2014 in Wiley Online Library (wileyonlinelibrary.com).

The human voice plays a key role in most social interactions not only because it conveys speech but also because it allows us to distinguish and recognize people. Even newborns are able to reliably differentiate their mother's voice from other voices [Beauchemin et al., 2011; DeCasper and Fifer, 1980; Kisilevsky et al., 2003]. Functional imaging studies have identified voice-sensitive regions along the superior temporal sulcus [STS, for a review, see Belin et al., 2004]. The STS seems to respond stronger to human vocalizations (both speech and non-speech) than to animal vocalization, other environmental or scrambled sounds not only in the mature [Belin et al., 2000, 2002; Fecteau et al.,

2004] but also in the developing human brain [Blasi et al., 2011; Grossmann et al., 2010]. Particularly, the right STS has been shown to be sensitive to speaker change [Belin and Zatorre, 2003] and preferentially process voice identity rather than the verbal content of speech [Belin and Zatorre, 2003; von Kriegstein et al., 2003; Von Kriegstein and Giraud, 2004]. Posterior regions of the STS have been reported to process speaker related acoustic differences in a speech signal [e.g., timbre, Andics et al., 2010; von Kriegstein, 2012; von Kriegstein et al., 2007, 2010], whereas anterior regions appear to be involved in identity processing of speech and non-speech signals [Andics et al., 2010; Belin and Zatorre, 2003; Imaizumi et al., 1997; Latinus et al., 2011; Nakamura et al., 2001; von Kriegstein et al., 2003; von Kriegstein and Giraud, 2004]. Furthermore, voice recognition has been reported to elicit activation in face-sensitive areas of the fusiform gyrus [von Kriegstein et al., 2005, 2008; von Kriegstein and Giraud, 2004, 2006], suggesting crossmodal interactions between face- and voice-processing areas for voice identification [Blank et al., 2011].

Blind individuals identify other people mainly through their voices. As for a number of other auditory functions [reviewed e.g., in Frasnelli et al., 2011; Merabet and Pascual-Leone, 2010; Pavani and Röder, 2012] improved processing [Föcker et al., 2012], learning [Föcker et al., 2012] and memory for voices [Bull et al., 1983; Röder and Neville, 2003] have been reported in blind compared with sighted adults. In addition, blind people have been observed to have a higher proficiency in discriminating voice prosodies [Klinge et al., 2010b]. Changes within auditory brain structure [intramodal plasticity, Röder and Neville, 2003], multisensory regions [De Volder et al., 1997; Röder et al., 1999], and a recruitment of visual cortices [crossmodal plasticity, Merabet and Pascual-Leone, 2010] have been suggested to mediate improved performance of the blind including voice processing [Gougoux et al., 2009].

Recent studies have provided evidence for some degree of a functional specialization within the visual cortex of the blind: While the processing of object identity [auditory object recognition: Amedi et al., 2010, tactile object recognition: Amedi et al., 2007; Pietrini et al., 2004, language recognition: Büchel et al., 1998a,b; Burton et al., 2002, 2003, 2006; Mahon et al., 2009; Noppeney et al., 2003; Reich et al., 2011; Röder et al., 2002] has consistently been found to activate the ventral part of the visual cortex; spatial processing [auditory localization: Collignon et al., 2007, 2011b; Gougoux et al., 2005; Renier et al., 2010; Voss et al., 2006; Weeks et al., 2000; auditory motion: Bedny et al., 2010; Poirier et al., 2006; Wolbers et al., 2011; tactile motion: Bonino et al., 2008; Matteau et al., 2010; Pfitz et al., 2009; Ricciardi et al., 2007] seems to recruit more dorsal parts of the occipital cortex. Thus, a functional specialization between a ventral stream and a dorsal stream, as observed in the visual modality [Ungerleider and Mishkin, 1982] and more recently within the auditory

modality [De Santis et al., 2007; Lomber and Malhotra, 2008] appears to be preserved in blind individuals' brain [Collignon et al., 2011b; Dormal and Collignon, 2011; Renier et al., ; Striem-Amit et al., 2012].

In this study, we addressed the question of whether voice identity processing is reorganized in blind people. Previous research has reported more activity in the STS for vocal vs. non-vocal sounds [Gougoux et al., 2009] and increased amygdala activation to fearful and angry voices [Klinge et al., 2010b] in congenitally blind individuals compared with sighted individuals. A recent ERP study [Föcker et al., 2012] has shown early ERP effects (between 100 and 160 ms) in a voice identity priming task in congenitally blind but not in sighted individuals. However, the precise neural sources of this activity remain unclear. The goal of this study was to gain more precise knowledge about the neural systems mediating voice identity processing in the blind and about the link between cross-modal reorganization and behavioral superiority of the blind. We first trained congenitally blind and matched sighted control participants to recognize unfamiliar voices in an extensive pre-experimental training and measured each participant's voice recognition skills. Thereafter, we used a priming paradigm, in which we manipulated whether two successively presented voices belonged to the same speaker or to different speakers. In the priming literature, it has been suggested that after the presentation of a prime subsequent processing is facilitated and requires less neural activity [Grill-Spector et al., 2006; Henson, 2003; Schacter and Buckner, 1998]. In line with this reasoning, functional magnetic resonance imaging (fMRI) studies on voice priming [Andics et al., 2010; Belin and Zatorre, 2003; Latinus et al., 2011] and face priming [Rotshtein et al., 2005; Winston et al., 2004] have shown that the BOLD signal declines with repeated presentations of identical stimuli. We therefore expected a decrease in activation in same-speaker (person-congruent) compared with different-speaker (person-incongruent) trials in regions that process voice identity, namely the STS and the fusiform gyrus. More specifically, we expected that activation in the fusiform gyrus would be modulated by speaker identity in blind but not in sighted participants.

METHODS

Participants

Twelve congenitally blind (six women, mean age: 36 years, age range: 23–48 years, nine right-handed, two ambidextrous) and 11 age and gender matched sighted individuals (mean age: 34 years, age range: 23–47 years, five female, 10 right-handed) participated in this study. The mean age did not differ between congenitally blind and sighted control participants ($t(21) = 0.51, P = 0.613$). Mean verbal intelligence scores [measured with the MWTB, German Mehrfach-Wortwahl-Test, Lehrl, 2005,

Figure 1.

Illustration of the experimental design. Two voice stimuli (disyllabic pseudowords) were successively presented. In 50% of the trials, S1 and S2 belonged to the same speaker (person-congruent voices); in the other 50%, S1 and S2 belonged to different

speakers (person-incongruent voices). Participants decided whether the S2 voice was from an old or from a young person. Additionally, participants had to detect deviant S1 stimuli (11.1% of all trials). ITI = inter-trial-interval.

applied in Braille to blind and in standard print to sighted participants] were above average in both groups and did not differ between groups (blind: 115 ± 3.6 (mean $\pm$ sem), sighted: 122 ± 4.1 , $t(21) = 1.35$, $P = 0.193$).

All blind participants were totally blind or did not have more than rudimentary sensitivity for brightness differences without any pattern vision. Blindness was due to peripheral reasons in all participants (retinopathy of prematurity ($n = 5$), retinoblastoma ($n = 2$), optic nerve atrophy ($n = 1$), perinatal hypoxia ($n = 1$), retina degeneration ($n = 1$), leber's congenital amaurosis ($n = 1$), and unknown peripheral defect ($n = 1$)). Sighted participants had normal or corrected to normal vision. All participants were German native speakers and reported normal hearing and no history of neurological illness. Hand preference was determined with the Edinburgh Handedness Inventory [Oldfield, 1971].

All participants were recruited from the local community or cities near the city of Hamburg and received monetary compensation for their participation. Written informed consent was given by each participant before the beginning of the experiment. This study was in accordance with the Declaration of Helsinki and approved by the Ethics committee of the medical association of Hamburg.

Experimental Design

Stimulus material

Stimulus material consisted of disyllabic German pseudowords spoken by 12 professional actors. The 12 actors were characterized by age and gender: three young women (mean age: 25 years, range: 23–27 years), three young men (mean age: 28 years, range: 26–29 years), three

old women (mean age: 63 years, range: 61–64 years), and three old men (mean age: 66 years, range: 56–79 years). Each talker's utterances were recorded in a sound-attenuated recording studio (Faculty of Media Technology at the Hamburg University of Applied Sciences) with a Neumann U87 microphone. Sound material was digitally sampled at 16 bit and offline equated for root mean square at 0.2 for presentation inside and at 0.025 for presentation outside the MR scanner. The mean duration of the auditory stimuli was 1044 ms (range: 676–1498 ms). To guarantee a smooth onset of the voice stimulus, a 50 ms period of silence was added before the actor's voicing began.

Procedure

Experiment. Within a S1-S2 paradigm, we presented two successive voice stimuli. Each trial began with a warning sound (550 Hz, duration = 100 ms). After an interval of 886–1889 ms (mean: 1217 ms), the first voice stimulus (S1) was presented and after an interstimulus interval (ISI) of 1150 ms the second voice stimulus (S2). The trial ended with the response of the participant, maximal 1000 ms after the offset of the second voice. Each trial was followed by a 4–12 s rest period (mean: 8 s, uniform distribution). In 50% of the trials, S1 and S2 belonged to the same speaker (person-congruent voices); in the other 50%, S1 and S2 belonged to different speakers (person-incongruent voices) (see Fig. 1). Participants decided whether the S2 voice was from an old or from a young person. An orthogonal task instead of an explicit speaker identity matching task was used in order to dissociate the effect of identity incongruency from response incongruency. Orthogonal tasks have been successfully used in other priming studies [Ellis et al., 1997; Föcker et al., 2011; Henson, 2003;

Noppeney et al., 2008]. Participants responded by pressing one of two buttons on a keypad with the index or the middle finger of the right hand. Response key assignments were counterbalanced across participants. For both conditions, 48 trials were presented resulting in a total number of 96 trials (standard trials). To guarantee attention to the S1 stimulus, 12 additional trials with deviant S1 stimuli were interspersed (deviant trials, 11.1% of all trials). Participants indicated the detection of a deviant stimulus by pressing the button which was assigned to the index finger. The experiment was presented in two sessions.

In standard trials, six pseudowords in which the first and second syllable were identical were presented (baba, dede, fafa, lolo, sasa, and wowo). In contrast, deviant S1 stimuli consisted of two different syllables (babu, dedu, fafi, lolu, and wowe). We used pseudowords in order to single out voice identity effects by minimizing possible confounds related with real words (e.g., semantic associations, valence, and familiarity). To avoid physically identical voice pairs in the person-congruent condition, different pseudowords were used as S1 and S2 in all conditions, e.g., “baba” as S1 and “dede” as S2. Stimuli were presented in pseudo-randomized order so that the same actor was never presented in two consecutive trials and deviant stimuli were separated by at least two standard stimuli. Overall, each actor was presented equally often as S1 and as S2. In person-incongruent trials, each speaker was paired once with a different speaker of the same age and gender, once with a different speaker of the same age but a different gender, once with a different speaker of a different age but the same gender, and once with a different speaker of a different age and a different gender. Consequently, 50% of person-incongruent trials (i.e., 25% of the total trials) were gender-congruent (S1 and S2 same gender) and 50% (i.e., 25% of the total trials) were gender-incongruent (S1 and S2 different gender). Similarly, 50% of person-incongruent trials were age-congruent (S1 and S2 same age) and 50% were age-incongruent (S1 and S2 different age). Note that age-congruent trials were also response-congruent (i.e., S1 primed response to S2) and age-incongruent trials response-incongruent (i.e., S1 did not prime response to S2). This procedure enabled us to disentangle the effect of voice identity from the effects of age, gender, and response.

Training. Before the experiment, participants were familiarized with all voice stimuli presented in standard trials in multiple extensive training sessions. Initially, all voice stimuli were introduced and associated with a disyllabic proper name for each actor. In each trial, participants listened to an auditorily presented name which was followed by one of six voice stimuli of the corresponding actor. Participants were instructed to memorize all name-voice associations. The main training consisted of two phases: a voice training phase and a voice matching phase. In the voice training phase, voice stimuli were presented and participants were asked to respond with the correct name of the actor. Feedback was provided after each response. Each training sequence con-

sisted of 36 voice stimuli, in which each actor was presented three times. This training phase ended as soon as the participant reached the criterion of 85% correct responses (31 out of 36 trials) in at least three consecutive training sequences. In the voice matching phase, voice stimuli were presented within a S1-S2 paradigm. One matching sequence consisted of 30 voice pairs of which 50% were person-congruent and 50% person-incongruent. In contrast to the main experiment, participants explicitly indicated whether the two voice stimuli belonged to the same or two different persons and received feedback after each response. Participants had to reach a criterion of 85% correct classifications in two successive blocks (26 out of 30 trials) to successfully terminate this training phase. Both, the voice training and the voice matching phase, were completed in each training session. At the end of the last training session, we tested whether participants were able to transfer their voice-specific knowledge to a novel set of stimuli. For each actor, eight new pseudowords (tete, gigi, nono, rara, babu, fafi, lolu, and wowe) were presented and participants were asked to provide the correct name. Participants did not receive any feedback for this task.

On the day of scanning, performance in the voice training and in the voice matching phase were assessed again outside the scanner. Furthermore, participants were familiarized with the experiment before scanning.

Data Acquisition

fMRI data were acquired on a 3 Tesla MR scanner (Siemens Magnetom Trio, Siemens, Erlangen, Germany) equipped with a 12 channel standard head coil. Thirty-six transversal slices (3 mm thickness, no gap) were acquired in each volume. A $T2^*$ -sensitive gradient echo-planar imaging (EPI) sequence was used (repetition time: 2.35 s, echo time: 30 ms, flip angle: 80° , field of view: 216×216 , matrix: 72×72). A 3D high-resolution ($1 \times 1 \times 1 \text{ mm}^3$ voxel size) T_1 -weighted structural MRI was acquired for each subject using a magnetization-prepared rapid gradient echo (MP-RAGE) sequence. Voice stimuli were presented via an MR-compatible electrodynamic headphone (MR confon GmbH, Magdeburg, Germany, <http://www.mr-confon.de>). Sound volume was adjusted to a comfortable level for each participant before the experiment. This ensured that stimuli were clearly audible for all participants. All participants were blindfolded throughout scanning. Task presentation and recording of behavioral responses were conducted with Presentation software (www.neurobs.com).

Data Analysis

Behavioral data

For each participant, the number of training sessions required to reach the learning criteria of 85% correctly recognized voices was determined. Furthermore, three performance measures were used to assess the voice recognition skills of each participant after the voice

training: (1) the voice recognition rate of new pseudowords (in %) at the end of the last training session (2) the voice recognition rate (in %) of familiar pseudowords on the day of scanning (pre-scanning voice recognition), and (3) the performance in the voice matching phase (in %) on the day of scanning (pre-scanning voice matching). The means of all four variables were calculated separately for blind and sighted participants and statistically compared using two-sample *t*-tests.

In the main experiment, reaction times (RTs) were analyzed relative to the onset of the S2 voice stimulus for standard stimuli and relative to the onset of the S1 voice stimulus for deviant stimuli. Trials with incorrect responses or too fast (before voice onset) or too slow responses (more than three standard deviations above a subject's mean in the respective condition) were excluded from all further analyses. For each participant, mean RTs and mean response accuracies were calculated separately for person-congruent trials, for person-incongruent trials, and for deviant trials. For standard trials, group differences in RTs and response accuracies were analyzed with an 2×2 ANOVA with the repeated measurement factor Voice Identity (person-congruent vs. person-incongruent) and the between-subject factor Group (blind vs. sighted). Mean response accuracies and RTs for deviant trials were statistically compared between groups using two-sample *t*-tests.

Two additional analyses within person-incongruent trials were performed in order to investigate the effects of age and gender priming. Mean RTs and response accuracies were calculated separately for age-congruent and age-incongruent and likewise for gender-congruent and gender-incongruent trials for each participant and then compared by paired *t* tests within each group.

fMRI data

Image processing and statistical analyses were performed with statistical parametric mapping (SPM8 software, Wellcome Department of Imaging Neuroscience, London, www.fil.ion.ucl.ac.uk/spm). The first five volumes of each session were discarded to allow for T1 saturation effects. Scans from each subject were realigned using the mean scan as a reference and corrected for susceptibility artifacts ("realign and unwarp"). The structural T1 image was coregistered to the mean functional image generated during realignment. The coregistered T1 image was then segmented into gray matter, white matter, and CSF using the unified segmentation approach provided with SPM8 [Ashburner and Friston, 2005]. This method has been shown to provide a better and more reliable matching of brains with structural abnormalities to a standard template and to result in greater sensitivity for functional activity than the commonly used alternatives, such as standard nonlinear approaches with cost-function masking [Crinion et al., 2007]. Functional images were subsequently spatially normalized to Montreal Neurologi-

cal Institute space using the normalization parameters obtained from the segmentation procedure, resampled to a voxel size of $2 \times 2 \times 2$ mm³, and spatially smoothed with a 8 mm full-width at half-maximum isotropic Gaussian kernel.

Statistical analysis was performed within a general linear model (GLM). We modeled person-congruent and person-incongruent trials as two separate event-related regressors (onset S2 stimulus, duration 0 s, only correct trials included) and convolved them with a hemodynamic response function. The statistical model further included deviant trials and trials with incorrect responses (errors) as regressors of no interest. Potential baseline drifts in time series were corrected by applying a high-pass frequency filter (128 s). To analyze age and gender priming effects within person-incongruent trials, we modified this model in that we modeled gender-congruent and gender-incongruent and respectively age-congruent and age-incongruent trials as two separate regressors in two additional models. For each participant, we created four contrast images: one to analyze voice identity priming (person-incongruent > person-congruent), one to analyze age priming within person-incongruent trials (age-incongruent > age-congruent), one to analyze gender priming within person-incongruent trials (gender-incongruent > gender-congruent), and one to analyze the mean activation in response to person-congruent and person-incongruent trials. The resulting four contrast images were then entered into a random effects group analysis. Between-group effects of voice identity priming and of the mean activation in response to person-congruent and person-incongruent trials were analyzed with two-sample *t*-tests. Within-group effects of voice identity, age, and gender priming were analyzed with one-sample *t*-tests. The pre-scanning voice recognition rate was included as a covariate in the within-group analyses in order to control for inter-individual differences in voice recognition skills on brain activation.

Activations at the group level were corrected for multiple comparisons using a family-wise error rate approach (FWE). For the occipital cortex and temporal lobe regions, we had a priori hypotheses and therefore limited our search volume to these regions. Correction for the occipital cortex was based on a mask for the occipital cortex taken from the Talairach Daemon database [Lancaster et al., 1997, 2000] created with the WFU PickAtlas version 3.0 [Maldjian et al., 2003, 2004]. For the fusiform gyrus and the STS, corrections were based on coordinates reported in previous studies. In detail, correction for the fusiform gyrus was based on a 14 mm radius sphere centered on *x*, *y*, *z*: ± 36 , -39 , -12 mm [Von Kriegstein and Giraud, 2004] and for the STS and adjacent cortices it was based on three 10 mm radius spheres centered on *x*, *y*, *z*: ± 63 , -34 , 7 for the posterior STS, on *x*, *y*, *z*: ± 63 , -7 , -14 for the middle STS and on *x*, *y*, *z*: ± 57 , 8 , -11 for the anterior STS [all from Blank et al., 2011]. All Talairach coordinates were transformed to MNI coordinates. For all other brain regions, correction was performed for all voxels.

TABLE I. For each group, the mean number of trainings sessions, the mean recognition rate of new pseudowords, the mean pre-scanning voice recognition rate, the mean pre-scanning voice matching rate, and the response accuracy for each trial type are shown

	Blind	Sighted
Training results		
Number of training sessions	1.7 (0.3)	2.6 (0.2)
Recognition of new pseudowords (%)	90.5 (1.5)	77.6 (1.7)
Pre-scanning voice recognition (%)	95.4 (1.3)	86.6 (2.3)
Pre-scanning voice matching (%)	97.8 (0.9)	95.8 (1.1)
Response accuracy in the experiment (%)		
Person-congruent trial	98.1 (0.5)	93.4 (1.6)
Person-incongruent trial	96.7 (0.8)	91.3 (1.7)
Deviant trial	98.6 (0.9)	100.0 (0.0)

The number in brackets indicates the standard error of the mean.

Activations in our regions of interest were correlated with the three voice recognition performance measures (pre-scanning voice recognition, pre-scanning voice matching, recognition of new pseudowords at the end of training) and the overall task performance (pooled over voice identity) in the main experiment. The individual activation was assessed by extracting the BOLD signal intensity of the peak voxel within the predefined spheres in each participant [using the rfxplot toolbox, <http://rfxplot.sourceforge.net/>, Gläscher, 2009].

Spatial references are reported in MNI standard space. For illustration purposes, statistical maps are thresholded at $P < 0.01$, uncorrected.

RESULTS

Behavioral Results

Training

Blind participants learned the voices within fewer training sessions than sighted control participants (Table I, $t(21) = 2.73$, $P = 0.013$). Six out of the 12 blind participants learned the voices within one training session while all sighted participants needed at least two. By contrast, one congenitally blind participant but five sighted participants needed more than two training sessions. Blind participants recognized a higher number of new pseudowords in the last training session (Table I, $t(21) = 5.78$, $P < 0.001$). Eleven blind but only two sighted participants recognized at least 85% of new pseudowords. On the day of scanning, blind participants recognized significantly more voices than sighted control participants (Table I, $t(21) = 3.39$, $P = 0.003$). Seven blind but only one sighted participants recognized all or all but one voice stimuli correctly. Blind did not differ from sighted control participants in the voice matching task (Table I, $t(21) = 1.39$, $P = 0.179$). Taken together, the voice recognition skills of blind participants

were superior to sighted control participants despite equal voice training procedures.

Main experiment

In the main experiment, response accuracies (Table I) were above 90% in all conditions. However, the overall response accuracy was significantly higher in blind participants than in sighted control participants (main effect of group, $F(1,21) = 10.17$, $P = 0.004$). RTs (Fig. 2) did not differ significantly between both groups ($F(1,21) = 1.28$, $P = 0.272$). Both groups responded more accurate ($F(1,21) = 7.11$, $P = 0.014$) and faster ($F(1,21) = 13.26$, $P = 0.002$) in person-congruent than in person-incongruent trials. There was no significant interaction, neither for response accuracies ($F(1,21) = 0.29$, $P = 0.593$) nor for RTs ($F(1,21) = 0.384$, $P = 0.542$).

The detection rate for S1 deviants was very high in both groups and did not differ between groups (Table I; $t(21) = 1.42$, $P = 0.171$). Blind participants responded faster to S1 deviants than sighted control participants (blind: mean response time of 1147 ms (22 ms); sighted: 1331 ms (67 ms); $t(21) = 2.69$, $P = 0.014$).

To control for gender and age priming effects, we directly compared response accuracies and reaction times between age-congruent and age-incongruent and between gender-congruent and gender-incongruent trials within person-incongruent trials (Supporting Information Table I). Both comparisons revealed no significant differences, neither in blind participants (age RTs: $t(11) = 1.07$, $P = 0.309$, age response accuracies: $t(11) = 1.10$, $P = 0.295$; gender RTs: $t(11) = 0.50$, $P = 0.630$, gender response accuracies: $t(11) = 0.583$, $P = 0.572$) nor in sighted control participants (age RTs: $t(10) = 1.79$, $P = 0.104$, age response accuracies: $t(10) = 0.28$, $P = 0.784$; gender RTs: $t(10) = 1.36$, $P = 0.202$, gender response accuracies: $t(10) = 1.12$, $P = 0.288$).

fMRI Data

Mean activation

The mean activation in response to person-congruent and person-incongruent trials was higher in the bilateral occipital cortex of blind participants than of sighted control participants (Fig. 3, peak coordinates x y z in mm, right peak: 44, -64, -2, $z = 4.62$, $P = 0.010$, left peak: -20, -82, 26, $z = 4.53$, $P = 0.014$, see Table II for whole brain results of this contrast).

Voice identity priming

In the right anterior fusiform gyrus, voice identity priming elicited a significantly higher activation in blind participants than in sighted control participants (Fig. 4, Table III, peak: 40, -36, -10, $z = 3.47$, $P = 0.043$). Within-group analyses showed activation of the right anterior fusiform gyrus in blind participants (peak: 40, -36, -6, $z = 3.54$, P

Figure 2.

Behavioral data. Mean response times in person-congruent and person-incongruent trials are shown for congenitally blind and sighted control participants. Response times were recorded from S2 onset onwards. Error bars indicate the standard error of the mean. Both groups responded significantly faster in person-congruent than in person-incongruent trials.

= 0.050), but not in sighted control participants ($P > 0.01$ uncorrected).

The right posterior STS showed a significant stronger voice identity priming effect in sighted control than in blind participants (Fig. 5, Table III; peak: 68 -30 12, $z = 3.29$, $P = 0.032$). Within-group analyses revealed no significant voice identity priming in the STS of blind participants ($P > 0.001$ uncorrected for the left posterior STS and $P > 0.01$ uncorrected for all other STS regions), but in the bilateral posterior STS of sighted control participants (left peak: -64, -28, 10, $z = 3.48$, $P = 0.036$; right peak: 62, -28, 8, $z = 3.51$, $P = 0.034$). In addition, voice identity priming effects were observed in the left precentral gyrus of sighted control participants (peak: -40, -2, 36, $z = 5.00$, $P = 0.049$ whole brain corrected).

Gender and age priming

Neither the STS nor the fusiform gyrus showed effects of gender or age priming (analyzed within person-incongruent trials) within blind or sighted control participants ($P > 0.01$ uncorrected).

Correlational analyses

No consistent relationships between any performance measure and brain activation in our regions of interest were observed. Note that the behavioral data showed little variance as participants performed at or near ceiling in all of these measures. For example, seven out of the 12 blind

participants recognized all or all but one voice stimuli in the pre-scanning voice recognition task, and 10 blind participants correctly matched all or all but one voice stimuli in the pre-scanning voice matching task.

DISCUSSION

The goal of this study was to identify the neural correlates of superior voice processing skills in congenitally blind humans. In congenitally blind but not in matched sighted control participants, the right anterior fusiform gyrus showed an increased BOLD signal in response to person-incongruent compared with person-congruent trials. Furthermore, voice identity priming was observed in the right posterior STS of sighted controls, but not in congenitally blind participants. Behaviorally, congenitally blind participants learned voices faster than sighted controls and displayed superior voice recognition skills after the training.

Our main finding implies the recruitment of the fusiform gyrus during auditory person identification in blind individuals. Crossmodal activations of ventral “visual” stream areas have been shown in a number of other higher-level cognitive tasks, e.g., recognition [auditory: Amedi et al., 2007; tactile: Amedi et al., 2010; Pietrini et al., 2004], verbal memory [Amedi et al., 2003] and semantic decisions [Noppeney et al., 2003]. It is still an open question whether occipital activation effectively facilitates nonvisual abilities in the blind or is merely an epiphenomenon [for a discussion see Pavani and Röder, 2012]. In favor of the first view are reports of positive correlations between behavioral measures and activations in striate [Amedi et al., 2003; Gougoux et al., 2005] and extrastriate areas [Gougoux et al., 2005] and disrupted verbal memory after TMS was applied over the occipital cortex [Amedi et al., 2004]. These findings suggest that the crossmodal activation of visual areas might mediate the blind’s superiority in a number of tasks [Collignon et al., 2011a] possibly including voice recognition [this study, Bull et al., 1983; Föcker et al., 2012; Röder and Neville, 2003] and voice learning [this study, Föcker et al., 2012] in the blind.

There is some evidence that the functional organization of extrastriate visual areas in sighted appears to be preserved in blind individuals [Renier et al., ; Voss and Zatorre, 2012]. For instance, the lateral-occipital complex (LOC), which responds to visual and tactile object shape in sighted [for a review see Lacey and Sathian, 2011], has been reported to respond during auditory shape processing in the blind [Amedi et al., 2007, 2010]. Similarly, separate areas of the ventral “visual” stream have been found to be activated during the tactile exploration of faces and objects [Pietrini et al., 2004] and for auditory living (e.g., faces, animals) and non-living stimuli [e.g., tools, houses, Mahon et al., 2009] in blind individuals. In contrast, the processing of spatial attributes of sounds has been observed to activate dorsal visual areas [Collignon et al.,

Figure 3.

Congenitally blind participants showed a stronger overall activation in the occipital cortex than sighted control participants. fMRI effects for the contrast blind > sighted (pooled over voice identity) are displayed. The mean percent signal change of the peak voxel is plotted for each group and separately for person-congruent and person-incongruent trials. Error bars indicate the standard error of the mean. L = left, R = right.

2011a,b; Renier et al., 2010]. One major difference between tasks activating dorsal or ventral visual areas might be their dependence on the retrieval of semantic information for the processing of stimuli. Voice recognition, or more generally, object recognition is accomplished through the interaction of perceptual and semantic processes, as it

TABLE II. Activation maxima for group comparisons, independent of voice identity effects (main effect of group)

Region	MNI coordinates			z-Value
	x	y	z	
Blind > Sighted				
R inferior occipital gyrus	46	-64	0	4.66 ^a
L superior occipital gyrus	-20	-82	26	4.53 ^a
L middle occipital gyrus	-40	-80	8	4.26
R temporale pole	36	20	-34	3.98
L fusiform gyrus	-32	-24	-26	3.60
R superior occipital gyrus	18	-84	36	3.51
L fusiform gyrus	-38	-50	-22	3.39
L temporale pole	-46	18	-26	3.38
L middle temporale gyrus	-62	-24	-8	3.27
Sighted > Blind				
L orbital superior frontal gyrus	18	26	-10	3.75
R Hippocampus	40	-20	-12	3.66
L Thalamus	-14	-14	20	3.43
L superior temporal gyrus	-62	-18	8	3.36

Coordinates are denoted by *x*, *y*, *z* in mm (MNI space) and indicate the peak voxel. Strength of activation is expressed in z-scores. Only activations with *P* < 0.001 uncorrected and 5 or more contiguous voxels are shown; L = left, R = right.

^a*P* < 0.1 whole brain corrected.

requires the association of the percept with stored semantic information (e.g., name) about the corresponding person or the corresponding object. Thus, our data is in line with previous studies [Collignon et al., 2011a,b; Renier et al., 2010; Striem-Amit et al., 2012] suggesting a functional segregation of ventral and dorsal cortical pathways in reorganized “visual” areas of congenitally blind humans. These reports suggest that the functional organization of extrastriate areas does not depend on visual experience [Voss and Zatorre, 2012]. This further implies that cortical structures may be primarily optimized for the operation that they perform rather than for a specific sensory input [Pascual-Leone and Hamilton, 2001]. Moreover, it has been proposed that cortical structures might switch their input modality as a consequence to missing sensory input but still maintain their original function [Lomber et al., 2010]. Consequently, crossmodal recruitment of deprived cortices should particularly exist for operations that are applied to inputs from different modalities [“supramodal functions,” Lomber et al., 2010]. Person recognition is a cognitive task which can be accomplished by using different modalities such as facial or vocal stimuli [Campanella and Belin, 2007; Schweinberger, 2013]. Therefore, one might speculate that the same areas of the fusiform gyrus that have been reported to be sensitive to face identity in sighted [Haxby et al., 2000; Rotshtein et al., 2005] may be sensitive to voice identity in the blind.

Interestingly, the reported activation in the anterior fusiform gyrus is in direct proximity to an area in which the recognition of voices has been reported to elicit activation in sighted individuals, but only for voices that had been associated with faces before the experiment [Von Kriegstein et al., 2005; Von Kriegstein and Giraud, 2004, 2006]. These data support the idea of a metamodal functional

Figure 4.

In the right fusiform gyrus, voice identity priming is higher in congenitally blind than in sighted control participants. fMRI effects are displayed for the two-way interaction (Blind > Sighted) $\times$ (person-incongruent > person-congruent). Activations are displayed on the MNI template. The mean percent sig-

nal change of the peak voxel is plotted for each group and separately for person-congruent and person-incongruent trials. Error bars indicate the standard error of the mean. L = left, R = right.

TABLE III. Activation maxima for group $\times$ voice identity interactions

Region	MNI coordinates			z-Value
	<i>x</i>	<i>y</i>	<i>z</i>	
(Blind > Sighted) × (person-incongruent > person-congruent)				
R fusiform gyrus	40	−36	−10	3.61 ^a
(Sighted > Blind) × (person-incongruent > person-congruent)				
L postcentral gyrus	−50	−8	48	4.18
L precentral gyrus	−38	−6	58	4.02
R precentral gyrus	52	0	46	3.80
R superior temporal gyrus/sulcus	68	−30	12	3.29 ^a

Coordinates are denoted by x , y , z in mm (MNI space) and indicate the peak voxel. Strength of activation is expressed in z-scores. Only activations with $P < 0.001$ uncorrected and 5 or more contiguous voxels are shown. L = left, R = right.

^a $P < 0.05$ small volume corrected.

organization of the brain in which cortical structures operate on input from different modalities [Pascual-Leone and Hamilton, 2001]. Moreover, these reports suggest that a crossmodal recruitment of the fusiform gyrus for voice identity processing occurs not only in adaptation to sensory loss but even in the typical developed brain and might allow for crossmodal recognition of individuals.

The pathways through which auditory information reaches the visual cortex are largely unknown. Changes in direct cortico-cortical connections have been discussed as one possible mechanism mediating crossmodal plasticity in the blind [Merabet and Pascual-Leone, 2010; Röder and Neville, 2003]. Evidence for the existence of direct cortico-cortical connections between different sensory cortical areas in healthy humans comes currently almost exclusively from animal studies [reviewed in Cappe et al., 2009, 2012]. In humans, different approaches to study connectivity have provided some indirect evidence for the existence of axonal connections between primary auditory and visual areas [Beer et al., 2011; Werner and Noppeney, 2010] and between face processing areas in the fusiform gyrus

Figure 5.

In the right STS, voice identity priming is higher in sighted control than in congenitally blind participants. fMRI effects are displayed for the two-way interaction (Sighted > Blind) $\times$ (person-incongruent > person-congruent). Activations are displayed on

the MNI template. The mean percent signal change of the peak voxel is plotted for each group and separately for person-congruent and person-incongruent trials. Error bars indicate the standard error of the mean. L = left, R = right.

and voice processing areas in the STS [Blank et al., 2011; von Kriegstein et al., 2005; von Kriegstein and Giraud, 2006]. Given that these direct connections between face processing areas in the fusiform gyrus and voice processing areas in the STS exist, one might speculate that congenital visual deprivation may induce a strengthening or expansion of these connections which in turn leads to a reallocation of voice identity processing from the STS to the fusiform gyrus. Consistent with this hypothesis, alterations in functional connectivity between primary sensory cortices have been demonstrated in the blind [Klinge et al., 2010a].

In contrast to the fusiform gyrus, the right posterior STS was less activated in congenitally blind than in matched sighted control participants. The posterior STS is thought to be involved in the analysis of acoustical changes of speech signals [Andics et al., 2010; von Kriegstein, 2012; Von Kriegstein et al., 2007, 2010] and is a well-established multisensory brain region [for a review see Beauchamp, 2005; Driver and Noesselt, 2008]. It has been suggested that visual deprivation might cause reorganization in the multisensory STS [Lewkowicz and Röder, 2012]. People with congenital bilateral cataracts, who had been blind for a few month after birth, did not show activation of the STS in response to visual stimuli during lipreading [Putzar et al., 2010] and failed to benefit from audiovisual presentation in speech recognition [Putzar et al., 2007]. These results suggest that the STS needs visual input to develop multisensory responsiveness [Lewkowicz and Röder, 2012]. One might speculate that the missing visual input is substituted by an enhanced responsiveness to sensory input from other modalities [Lewkowicz and Röder, 2012]. This hypothesis is supported by a previous study, in which the authors demonstrated that vocal compared with nonvocal stimuli elicited larger activity in the STS in con-

genitally blind compared with sighted and late blind individuals [Gougoux et al., 2009]. These data suggest that intramodal plasticity could possibly increase the efficiency of perceptual processing of voices in the blind. In contrast to a pure perceptual analysis of voices, recognizing voice identities involves multimodal processing in sighted individuals, i.e., the association of visual, auditory, and semantic identity information. The lack of visual input during development might result in a reduced engagement of multisensory areas of the STS during voice identity processing in the blind. Taking into account the voice-identity related activation in the anterior fusiform gyrus and the evidence for direct pathways between STS and the anterior fusiform gyrus [Blank et al., 2011], one might even hypothesize that voice identity processing is reallocated from the STS to the anterior fusiform gyrus in congenitally blind individuals.

In sum, this study suggests a functional adaption of the person identification system following congenital blindness. Specifically, we report a crossmodal recruitment of the fusiform gyrus during the processing of voice identity. A recent ERP study with the same stimuli, paradigm, and a subsample of the same congenitally blind participants [Föcker et al., 2012] suggested that a reorganization of the person identification system appears to affect early perceptual processes starting around 100 ms poststimulus onset. Moreover, studies with sighted adults have suggested direct connections between voice processing areas in the STS and in the fusiform gyrus [Blank et al., 2011; Von Kriegstein et al., 2005; von Kriegstein and Giraud, 2006]. One might speculate that the lack of visual input results in a strengthening of these connections, which possibly permits a reallocation of voice identity processing from the STS to the fusiform gyrus in congenitally blind individuals.

ACKNOWLEDGMENTS

We thank Katrin Wendt, Kathrin Müller, and Corinna Klinge with their support acquiring the fMRI data and Jürgen Finsterbusch for setting up the fMRI sequence. We are grateful to Boris Schlaack for his support to create the stimulus material and to Ulrike Adam, Kirstin Grewenig, and Florence Kroll for helping to record the stimulus material supervised by Prof. Dr. Eva Wilk. We thank the “Blinden-und Sehbehindertenverein Hamburg, e.V.”, the “Dialogue of the Dark” in Hamburg, and the “Tandem-Club Weisse Speicher Hamburg e.V.” for their help in recruiting blind participants.

REFERENCES

- Amedi A, Raz N, Pianka P, Malach R, Zohary E (2003): Early “visual” cortex activation correlates with superior verbal memory performance in the blind. *Nat Neurosci* 6:758–766.
- Amedi A, Floel A, Knecht S, Zohary E, Cohen LG (2004): Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nat Neurosci* 7:1266–1270.
- Amedi A, Stern WM, Camprodon JA, Bermpohl F, Merabet L, Rotman S, Hemond C, Meijer P, Pascual-Leone A (2007): Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat Neurosci* 10:687–689.
- Amedi A, Raz N, Azulay H, Malach R, Zohary E (2010): Cortical activity during tactile exploration of objects in blind and sighted humans. *Restor Neurol Neurosci* 28:143–156.
- Andics A, McQueen JM, Petersson KM, Gál V, Rudas G, Vidnyánszky Z (2010): Neural mechanisms for voice recognition. *NeuroImage* 52:1528–1540.
- Ashburner J, Friston KJ (2005): Unified segmentation. *NeuroImage* 26:839–851.
- Beauchamp MS (2005): See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Curr Opin Neurobiol* 15:145–153.
- Beauchemin M, González-Frankenberger B, Tremblay J, Vannasing P, Martínez-Montes E, Belin P, Béland R, Francoeur D, Carceller A-M, Wallois F, Lassonde M (2011): Mother and stranger: An electrophysiological study of voice processing in newborns. *Cereb Cortex* 21:1705–1711.
- Bedny M, Konkle T, Pelphrey K, Saxe R, Pascual-Leone A (2010): Sensitive period for a multimodal response in human visual motion area MT/MST. *Curr Biol* 20:1900–1906.
- Beer AL, Plank T, Greenlee MW (2011): Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Exp Brain Res* 213:299–308.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000): Voice-selective areas in human auditory cortex. *Nature* 403:309–312.
- Belin P, Zatorre RJ, Ahad P (2002): Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res* 13:17–26.
- Belin P, Fecteau S, Bédard C (2004): Thinking the voice: Neural correlates of voice perception. *Trends Cogn Sci* 8:129–135.
- Belin P, Zatorre RJ (2003): Adaptation to speaker’s voice in right anterior temporal lobe. *Neuroreport* 14:2105–2109.
- Blank H, Anwender A, von Kriegstein K (2011): Direct structural connections between voice- and face-recognition areas. *J Neurosci* 31:12906–12915.
- Blasi A, Mercure E, Lloyd-Fox S, Thomson A, Brammer M, Sauter D, Deeley Q, Barker GJ, Renvall V, Deoni S, Gasston D, Williams SCR, Johnson MH, Simmons A, Murphy DGM (2011): Early specialization for voice and emotion processing in the infant brain. *Curr Biol* 21:1220–1224.
- Bonino D, Ricciardi E, Sani L, Gentili C, Vanello N, Guazzelli M, Vecchi T, Pietrini P (2008): Tactile spatial working memory activates the dorsal extrastriate cortical pathway in congenitally blind individuals. *Arch Ital Biol* 146:133–146.
- Büchel C, Price C, Frackowiak RS, Friston K (1998a): Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain* 121 (Part 3):409–419.
- Büchel C, Price C, Friston K (1998b): A multimodal language region in the ventral visual pathway. *Nature* 394:274–277.
- Bull R, Rathborn H, Clifford BR (1983): The voice-recognition accuracy of blind listeners. *Perception* 12:223–226.
- Burton H, Snyder AZ, Diamond JB, Raichle ME (2002): Adaptive changes in early and late blind: A fMRI study of verb generation to heard nouns. *J Neurophysiol* 88:3359–3371.
- Burton H, Diamond JB, McDermott KB (2003): Dissociating cortical regions activated by semantic and phonological tasks: A fMRI study in blind and sighted people. *J Neurophysiol* 90:1965–1982.
- Burton H, McLaren DG, Sinclair RJ (2006): Reading embossed capital letters: an fMRI study in blind and sighted individuals. *Hum Brain Mapp* 27:325–339.
- Campanella S, Belin P (2007): Integrating face and voice in person perception. *Trends Cogn Sci* 11:535–543.
- Cappe C, Rouiller EM, Barone P (2009): Multisensory anatomical pathways. *Hear Res* 258:28–36.
- Cappe C, Rouiller EM, Barone P (2012): Cortical and thalamic pathways for multisensory and sensorimotor interplay. In: Murray MM, Wallace MT, editors. *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press. *Frontiers in Neuroscience*. <http://www.ncbi.nlm.nih.gov/books/NBK92866/>.
- Collignon O, Lassonde M, Lepore F, Bastien D, Veraart C (2007): Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cereb Cortex* 17:457–465.
- Collignon O, Champoux F, Voss P, Lepore F (2011a): Sensory rehabilitation in the plastic brain. *Prog Brain Res* 191:211–231.
- Collignon O, Vandewalle G, Voss P, Albouy G, Charbonneau G, Lassonde M, Lepore F (2011b): Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *Proc Natl Acad Sci USA* 108:4435–4440.
- Crinin J, Ashburner J, Leff A, Brett M, Price C, Friston K (2007): Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *NeuroImage* 37:866–875.
- De Santis L, Spierer L, Clarke S, Murray MM (2007): Getting in touch: Segregated somatosensory what and where pathways in humans revealed by electrical neuroimaging. *NeuroImage* 37:890–903.
- De Volder AG, Bol A, Blin J, Robert A, Arno P, Grandin C, Michel C, Veraart C (1997): Brain energy metabolism in early blind subjects: Neural activity in the visual cortex. *Brain Res* 750:235–244.
- DeCasper AJ, Fifer WP (1980): Of human bonding: Newborns prefer their mothers’ voices. *Science* 208:1174–1176.
- Dormal G, Collignon O (2011): Functional selectivity in sensory-deprived cortices. *J Neurophysiol* 105:2627–2630.
- Driver J, Noesselt T (2008): Multisensory interplay reveals cross-modal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron* 57:11–23.

- Ellis HD, Jones DM, Mosdell N (1997): Intra- and inter-modal repetition priming of familiar faces and voices. *Br J Psychol* 88 (Part 1):143–156.
- Fecteau S, Armony JL, Joannette Y, Belin P (2004): Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage* 23:840–848.
- Föcker J, Hölzig C, Best A, Röder B (2011): Crossmodal interaction of facial and vocal person identity information: An event-related potential study. *Brain Res* 1385:229–245.
- Föcker J, Best A, Hölzig C, Röder B (2012): The superiority in voice processing of the blind arises from neural plasticity at sensory processing stages. *Neuropsychologia* 50:2056–2067.
- Frasnelli J, Collignon O, Voss P, Lepore F (2011): Crossmodal plasticity in sensory loss. *Prog Brain Res* 191:233–249.
- Gläscher J (2009): Visualization of group inference data in functional neuroimaging. *Neuroinformatics* 7:73–82.
- Gougoux F, Zatorre RJ, Lassonde M, Voss P, Lepore F (2005): A functional neuroimaging study of sound localization: Visual cortex activity predicts performance in early-blind individuals. *PLOS Biol* 3:e27–e27.
- Gougoux F, Belin P, Voss P, Lepore F, Lassonde M, Zatorre RJ (2009): Voice perception in blind persons: A functional magnetic resonance imaging study. *Neuropsychologia* 47:2967–2974.
- Grill-Spector K, Henson R, Martin A (2006): Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23.
- Grossmann T, Oberecker R, Koch SP, Friederici AD (2010): The developmental origins of voice processing in the human brain. *Neuron* 65:852–858.
- Haxby JV, Hoffman EA, Gobbini MI (2000): The distributed human neural system for face perception. *Trends Cogn Sci* 4: 223–233.
- Henson RNA (2003): Neuroimaging studies of priming. *Prog Neurobiol* 70:53–81.
- Imaizumi S, Mori K, Kiritani S, Kawashima R, Sugiura M, Fukuda H, Itoh K, Kato T, Nakamura A, Hatano K, Kojima S, Nakamura K (1997): Vocal identification of speaker and emotion activates different brain regions. *Neuroreport* 8:2809–2812.
- Kisilevsky BS, Hains SMJ, Lee K, Xie X, Huang H, Ye HH, Zhang K, Wang Z (2003): Effects of experience on fetal voice recognition. *Psychol Sci* 14:220–224.
- Klinge C, Eippert F, Röder B, Büchel C (2010a): Corticocortical connections mediate primary visual cortex responses to auditory stimulation in the blind. *J Neurosci* 30:12798–12805.
- Klinge C, Röder B, Büchel C (2010b): Increased amygdala activation to emotional auditory stimuli in the blind. *Brain* 133:1729–1736.
- Lacey S, Sathian K (2011): Multisensory object representation: Insights from studies of vision and touch. *Prog Brain Res* 191: 165–176.
- Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AC, Toga AW, Mazziotta JC (1997): Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Hum Brain Mapp* 5:238–242.
- Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT (2000): Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 10:120–131.
- Latinus M, Crabbe F, Belin P (2011): Learning-induced changes in the cerebral processing of voice identity. *Cereb Cortex* 21:2820–2828.
- Lehrl S (2005): Manual zum MWT-B: [Mehrfachwahl-Wortschatz-Intelligenztest]. Balingen: Spitta-Verl.
- Lewkowicz DJ, Röder B (2012): Development of multisensory processes and the role of early experience. In: Stein BE, editor. *The New Handbook of Multisensory Processes*. Cambridge: MIT Press. pp 607–626.
- Lomber SG, Malhotra S (2008): Double dissociation of “what” and “where” processing in auditory cortex. *Nat Neurosci* 11:609–616.
- Lomber SG, Meredith MA, Kral A (2010): Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat Neurosci* 13:1421–1427.
- Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A (2009): Category-specific organization in the human brain does not require visual experience. *Neuron* 63:397–405.
- Maldjian JA, Laurienti PJ, Burdette JB, Kraft RA (2003): An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* 19:1233–1239.
- Maldjian JA, Laurienti PJ, Burdette JH (2004): Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage* 21:450–455.
- Matteau I, Kupers R, Ricciardi E, Pietrini P, Ptito M (2010): Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Res Bull* 82:264–270.
- Merabet LB, Pascual-Leone A (2010): Neural reorganization following sensory loss: The opportunity of change. *Nat Rev Neurosci* 11:44–52.
- Nakamura K, Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, Ito K, Kojima S (2001): Neural substrates for recognition of familiar voices: A PET study. *Neuropsychologia* 39:1047–1054.
- Noppeney U, Friston KJ, Price CJ (2003): Effects of visual deprivation on the organization of the semantic system. *Brain* 126: 1620–1627.
- Noppeney U, Josephs O, Hocking J, Price CJ, Friston KJ (2008): The effect of prior visual information on recognition of speech and sounds. *Cereb Cortex* 18:598–609.
- Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Pascual-Leone A, Hamilton R (2001): The metamodal organization of the brain. *Prog Brain Res* 134:427–445.
- Pavani F, Röder B (2012): Crossmodal plasticity as a consequence of sensory loss: Insights from blindness and deafness. In: Stein BE, editor. *The New Handbook of Multisensory Processes*. Cambridge: MIT Press. pp 737–759.
- Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu WHC, Cohen L, Guazzelli M, Haxby JV (2004): Beyond sensory images: Object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658–5663.
- Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG (2006): Auditory motion perception activates visual motion areas in early blind subjects. *NeuroImage* 31:279–285.
- Ptito M, Matteau I, Gjedde A, Kupers R (2009): Recruitment of the middle temporal area by tactile motion in congenital blindness. *Neuroreport* 20:543–547.
- Putzar L, Goerendt I, Lange K, Rosler F, Roder B (2007): Early visual deprivation impairs multisensory interactions in humans. *Nat Neurosci* 10:1243–1245.
- Putzar L, Goerendt I, Heed T, Richard G, Büchel C, Röder B (2010): The neural basis of lip-reading capabilities is altered by early visual deprivation. *Neuropsychologia* 48:2158–2166.

- Reich L, Szwed M, Cohen L, Amedi A (2011): A ventral visual stream reading center independent of visual experience. *Curr Biol* 21:363–368.
- Renier LA, Anurova I, De Volder AG, Carlson S, VanMeter J, Rauschecker JP (2010): Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind. *Neuron* 68:138–148.
- Renier L, De Volder AG, Rauschecker JP (in press): Cortical plasticity and preserved function in early blindness. *Neurosci Biobehav Rev*. <http://dx.doi.org/10.1016/j.neubiorev.2013.01.025> (in press).
- Ricciardi E, Vanello N, Sani L, Gentili C, Scilingo EP, Landini L, Guazzelli M, Bichi A, Haxby JV, Pietrini P (2007): The effect of visual experience on the development of functional architecture in hMT+. *Cereb Cortex* 17:2933–2939.
- Röder B, Neville H (2003): Developmental functional plasticity. In: Boller F, Grafman J, editors. *Plasticity and Rehabilitation. Handbook of Neuropsychology*. Amsterdam: Elsevier. pp 231–270.
- Röder B, Rösler F, Neville HJ (1999): Effects of interstimulus interval on auditory event-related potentials in congenitally blind and normally sighted humans. *Neurosci Lett* 264:53–56.
- Röder B, Stock O, Bien S, Neville H, Rösler F (2002): Speech processing activates visual cortex in congenitally blind humans. *Eur J Neurosci* 16:930–936.
- Rotshtein P, Henson RNA, Treves A, Driver J, Dolan RJ (2005): Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8:107–113.
- Schacter DL, Buckner RL (1998): Priming and the brain. *Neuron* 20:185–195.
- Schweinberger SR (2013): Audiovisual integration in speaker identification. In: Belin P, Campanella S, Ethofer T, editors. *Integrating Face and Voice in Person Perception*. New York: Springer. pp 119–134.
- Striem-Amit E, Dakwar O, Reich L, Amedi A (2012): The large-scale organization of “visual” streams emerges without visual experience. *Cereb Cortex* 22:1698–1709.
- Ungerleider LG, Mishkin M (1982): Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield, RJW, editors. *Analysis of visual behavior*. Cambridge, MA: The MIT Press. pp 549–586.
- Von Kriegstein K (2012): A multisensory perspective on human auditory communication. In: Murray MM, Wallace MT, editors. *The Neural Bases of Multisensory Processes*. Boca Raton (FL): Frontiers in Neuroscience. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22593871>.
- Von Kriegstein K, Eger E, Kleinschmidt A, Giraud AL (2003): Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Res Cogn Brain Res* 17: 48–55.
- Von Kriegstein K, Kleinschmidt A, Sterzer P, Giraud A-L (2005): Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci* 17:367–376.
- Von Kriegstein K, Smith DRR, Patterson RD, Ives DT, Griffiths TD (2007): Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Curr Biol* 17: 1123–1128.
- Von Kriegstein K, Dogan O, Grüter M, Giraud A-L, Kell CA, Grüter T, Kleinschmidt A, Kiebel SJ (2008): Simulation of talking faces in the human brain improves auditory speech recognition. *Proc Natl Acad Sci USA* 105:6747–6752.
- Von Kriegstein K, Smith DRR, Patterson RD, Kiebel SJ, Griffiths TD (2010): How the human brain recognizes speech in the context of changing speakers. *J Neurosci* 30:629–638.
- Von Kriegstein K, Giraud AL (2004): Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage* 22:948–955.
- Von Kriegstein K, Giraud AL (2006): Implicit multisensory associations influence voice recognition. *PLOS Biol* 4:e326–e326.
- Voss P, Gougoux F, Lassonde M, Zatorre RJ, Lepore F (2006): A positron emission tomography study during auditory localization by late-onset blind individuals. *Neuroreport* 17:383–388.
- Voss P, Zatorre RJ (2012): Organization and reorganization of sensory-deprived cortex. *Curr Biol* 22:R168–R173.
- Weeks R, Horwitz B, Aziz-Sultan A, Tian B, Wessinger CM, Cohen LG, Hallett M, Rauschecker JP (2000): A positron emission tomographic study of auditory localization in the congenitally blind. *J Neurosci* 20:2664–2672.
- Werner S, Noppeney U (2010): Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J Neurosci* 30:2662–2675.
- Winston JS, Henson RN, Fine-Goulden MR, Dolan RJ (2004): fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *J Neurophysiol* 92: 1830–1839.
- Wolbers T, Klatzky RL, Loomis JM, Wutte MG, Giudice NA (2011): Modality-independent coding of spatial layout in the human brain. *Curr Biol* 21:984–989.