# Cortical Network Differences in the Sighted Versus Early Blind for Recognition of Human-Produced Action Sounds

James W. Lewis,[1,2]* Chris Frum,[1,2] Julie A. Brefczynski-Lewis,[1,2,3]
William J. Talkington,[1,2] Nathan A. Walker,[1,2] Kristina M. Rapuano,[1,2]
and Amanda L. Kovach[1,2]

[1]Department of Physiology and Pharmacology, West Virginia University, Morgantown, West Virginia
[2]Center for Advanced Imaging, West Virginia University, Morgantown, West Virginia
[3]Department of Radiology, West Virginia University, Morgantown, West Virginia

◆ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ◆

**Abstract:** Both sighted and blind individuals can readily interpret meaning behind everyday real-world sounds. In sighted listeners, we previously reported that regions along the bilateral posterior superior temporal sulci (pSTS) and middle temporal gyri (pMTG) are preferentially activated when presented with recognizable action sounds. These regions have generally been hypothesized to represent primary loci for complex motion processing, including visual biological motion processing and audio–visual integration. However, it remained unclear whether, or to what degree, life-long visual experience might impact functions related to hearing perception or memory of sound-source actions. Using functional magnetic resonance imaging (fMRI), we compared brain regions activated in congenitally blind versus sighted listeners in response to hearing a wide range of recognizable human-produced action sounds (excluding vocalizations) versus unrecognized, backward-played versions of those sounds. Here, we show that recognized human action sounds commonly evoked activity in both groups along most of the left pSTS/pMTG complex, though with relatively greater activity in the right pSTS/pMTG by the blind group. These results indicate that portions of the postero-lateral temporal cortices contain domain-specific hubs for biological and/or complex motion processing independent of sensory-modality experience. Contrasting the two groups, the sighted listeners preferentially activated bilateral parietal plus medial and lateral frontal networks, whereas the blind listeners preferentially activated left anterior insula plus bilateral anterior calcarine and medial occipital regions, including what would otherwise have been visual-related cortex. These global-level network differences suggest that blind and sighted listeners may preferentially use different memory retrieval strategies when hearing and attempting to recognize action sounds. *Hum Brain Mapp 32:2241–2255, 2011.* © **2011 Wiley Periodicals, Inc.**

**Key words:** hearing perception; episodic memory; mirror-neuron systems; cortical plasticity; fMRI

◆ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ◆

## INTRODUCTION

Is the human cortical organization for sound recognition influenced by life-long visual experience? In young adult sighted participants, we previously identified left-lateralized networks of "high-level" cortical regions, well beyond early and intermediate processing stages of auditory cortex proper, which were preferentially activated when everyday, nonverbal real-world sounds were judged as accurately recognized—in contrast to hearing backward-played versions of those same sounds that were unrecognized [Lewis et al., 2004]. These networks appeared to include at least two multisensory-related systems. One system involved left inferior parietal cortex, which was later shown to have a role in audio-motor associations that depended on handedness [Lewis et al., 2005, 2006] and to overlap mirror-neuron systems in both sighted [Rizzolatti and Craighero, 2004; Rizzolatti et al., 1996] and blind listeners [Ricciardi et al., 2009]. Another system involved the left and right postero-lateral temporal regions, which seemed consistent with functions related to audio–visual motion integration or associations. This included the posterior superior temporal sulci (pSTS) and posterior middle temporal gyri (pMTG), herein referred to as the pSTS/pMTG complexes.

Visual response properties of the pSTS/pMTG complexes have been well studied. The postero-lateral temporal regions are prominently activated when viewing videos of human (conspecific) biological motions [Johansson, 1973], such as human face or limb actions [Beauchamp et al., 2002; Calvert and Brammer, 1999; Calvert and Campbell, 2003; Grossman and Blake, 2002; Grossman et al., 2000; Kable et al., 2005; Puce and Perrett, 2003; Puce et al., 1998] and point light displays of human actions [Safford et al., 2010]. In congenitally deaf individuals, cortex overlapping or near the pSTS/pMTG regions show a greater expanse of activation to visual motion processing when monitoring peripheral locations of motion flow fields [Bavelier et al., 2000], attesting to a prominent role in visual motion processing.

Hearing perception studies have also revealed a prominent role of the pSTS/pMTG complexes in response to human action sounds (i.e., conspecific action sounds, excluding vocalizations), which may be regarded as one distinct category of action sound. This includes, for example, assessing footstep movements [Bidet-Caulet et al., 2005], discriminating hand-tool action sounds from animal vocalizations [Lewis et al., 2005, 2006], and hearing hand-executed action sounds relative to environmental sounds [Gazzola et al., 2006; Ricciardi et al., 2009]. Furthermore, the pSTS/pMTG complexes show category-preferential activation for human-produced action sounds when contrasted with action sounds produced by nonhuman animals or nonliving sound-sources such as automated machinery and the natural environment [Engel et al., 2009; Lewis et al., in press].

The pSTS/pMTG regions, anatomically situated midway between early auditory and early visual processing corti-ces, are also commonly implicated in functions related to audio–visual integration, showing enhanced activations to audio–visual inputs derived from natural scenes, such as talking faces, lip reading, or observing hand tools in use [Beauchamp et al., 2004a,b; Calvert et al., 1999, 2000; Kreifelts et al., 2007; Olson et al., 2002; Robins et al., 2009; Stevenson and James, 2009; Taylor et al., 2006, 2009]. Such findings have provided strong support for the proposal that the postero-lateral temporal regions are primary loci for complex natural motion processing [for reviews see Lewis, 2010; Martin, 2007]. Postscanning interviews from our earlier study [Lewis et al., 2004] revealed that some of the sighted participants would "visualize" the sound-source upon hearing it (e.g., visualizing a person's hands typing on a keyboard). One possibility was that visual associations or "visual imagery" might have been evoked by the action sounds when attempting to recognize them, thereby explaining the activation of the bilateral pSTS/pMTG complexes. Thus, one hypothesis we sought to test was that if the pSTS/pMTG regions are involved in some form of visual imagery of sound-source actions, then there should be no activation of these regions in listeners who have never had visual experience (i.e., congenitally blind).

In addition to probing the functions of the pSTS/pMTG complexes, we additionally sought to identify differences in global network activations (including occipital cortices) that might be differentially recruited by congenitally blind listeners. A recent study using hand-executed action sounds reported activation to more familiar sounds near middle temporal regions in both sighted and blind listeners [Ricciardi et al., 2009]. However, their study focused on fronto-parietal networks associated with mirror-neuron systems or their analogues in blind listeners. Thus, it remained unclear what *global* network differences might exist between sighted and blind listeners for representing acoustic knowledge and other processes that are related to attaining a sense of recognition of real-world sounds.

Congenitally blind listeners are reported to show better memory for environmental sounds after physical or semantic encoding [Roder and Rosler, 2003]. Numerous studies involving early blind individuals indicate that occipital cortices, in what would otherwise have been predominantly visual-related cortices, become recruited to subserve a variety of other functions and possibly confer compensatory changes in sensory and cognitive processing. For instance, occipital cortices of blind individuals are known to adapt to facilitate linguistic functions [Burton, 2003; Burton and McLaren, 2006; Burton et al., 2002a,b; Hamilton and Pascual-Leone, 1998; Hamilton et al., 2000; Sadato et al., 1996, 2002], verbal working memory skills [Amedi et al., 2003; Roder et al., 2002], tactile object recognition [Pietrini et al., 2004], object shape processing [Amedi et al., 2007], sound localization [Gougoux et al., 2005], and motion processing of artificial (tonal) acoustic signals [Poirier et al., 2006]. Thus, a second hypothesis we sought to test was that the life-long audio–visual-motor experiences of sighted listeners, relative to the audio-motor

experiences of blind listeners, will lead to large-scale network differences in cortical organization for representing knowledge, or memory, of real-world human action sounds.

## MATERIALS AND METHODS

### Participants

Native English speakers with self reported normal hearing and who were neurologically normal (excepting visual function) participated in the imaging study. Ten congenitally blind volunteer participants were included (average age of 54; ranging from 38 to 64, seven female, one left-handed), together with 14 age-matched sighted control (SC) participants (average age 54; ranging 37 to 63, seven female; one left-handed). We obtained a neurologic history for each blind participant using a standard questionnaire. The cause of blindness for nine participants was retinopathy of prematurity (formerly known as retrolental fibroplasia), and one participant had an unknown cause of blindness. All reported that they had lost sight (were told they had lost sight) at or within a few months after birth, and herein referred to as early blind (EB). Some participants reported that they could remember seeing shadows at early ages (<6 years) but could never discriminate visual objects. All EB participants were proficient Braille readers. We assessed handedness with a modified Edinburgh handedness inventory [Oldfield, 1971], based on 12 questions, substituting for blind individuals the question "preferred hand for writing" with "preferred hand for Braille reading when required to use one hand." Informed consent was obtained for all participants following guidelines approved by the West Virginia University Institutional Review Board, and all were paid an hourly stipend.

### Sound Stimuli and Delivery

Sound stimuli included 105 real-world sounds described in our earlier study [Lewis et al., 2004], which were compiled from professional CD collections (Sound Ideas, Richmond Hill, Ontario, Canada) and from various web sites (44.1 kHz, 16-bit, monophonic). The sounds were trimmed to ~2 s duration (1.1–2.5 s range) and were temporally reversed to create "backward" renditions of those sounds (Cool Edit Pro, Syntrillium Software, now owned by Adobe). The backward-played sounds were chosen as a control condition because they were typically judged to be unrecognizable, yet were precisely matched for many low-level acoustic signal features, including overall intensity, duration, spectral content, spectral variation, and acoustic complexity. However, the backward sounds did necessarily differ in their temporal envelopes, having different attacks and offsets. Sound stimuli were delivered using a Windows PC computer, using Presentation software (version 11.1, Neurobehavioral Systems) via a sound mixer

and MR compatible electrostatic ear buds (STAX SRS-005 Earspeaker system; Stax, Gardena, CA), worn under sound attenuating ear muffs. Stimulus loudness was set to a comfortable level for each participant, typically 80–83 dB C-weighted in each ear (Brüel & Kjær 2239a sound meter), as assessed at the time of scanning.

During each fMRI scan, subjects indicated by three alternative forced choice (3AFC) right hand button press whether they (1) could recognize or identify the sound (i.e., verbalize, describe, imagine, or have a high degree of certainty about what the likely sound-source was), (2) were unsure, or (3) knew that they did not recognize the sound. Each participant underwent a brief training session just before scanning wherein several backward- and forward-played practice sounds were presented: If the participant could verbally identify it then they were affirmed of this and instructed to press button no. 1. If he or she had no idea what the sound-source was or could only hazard a guess, then they were instructed to press button no. 3. Participant's were instructed that they could use a second button press (button no. 2) for instances where they were hesitant about identifying the sound or felt that if given more time or a second presentation that they might be able to guess what it was. Each backward-played sound stimulus was presented before the corresponding forward-played version within a scanning run to avoid potential priming effects; participants were not informed that the "modified" sounds were simply played backwards. Overall, this paradigm relied on the novelty of having participants hearing each unique sound out of context for the first time, and their indication of whether or not the sound evoked a sense of a recognizable action event. The sighted individuals were asked to keep their eyes closed throughout all of the functional scans.

### Magnetic Resonance Imaging and Data Analyses

Scanning was conducted on a 3 Tesla General Electric Horizon HD MRI scanner using a quadrature bird-cage head coil. For the main paradigm, we acquired whole-head, spiral in and out imaging of blood-oxygenated level dependent (BOLD) signals [Glover and Law, 2001], using a clustered-acquisition fMRI design which allowed sound stimuli to be presented during scanner silence [Edmister et al., 1999; Hall et al., 1999]. A sound or a silent event was presented every 9.3 s, with each event being triggered by the MRI scanner. Button responses and reaction times relative to sound onset were collected during scanning. BOLD signals were collected 6.5 s after sound or silent event onset (28 axial brain slices, $1.875 \times 1.875 \times 4.00$ mm$^3$ spatial resolution, TE = 36 ms, OPTR = 2.3 s volume acquisition, FOV = 24 mm). This volume covered the entire brain for all subjects. Whole brain T1-weighted anatomical MR images were collected using a spoiled GRASS pulse sequence (SPGR, 1.2 mm slices with $0.9375 \times 0.9375$ mm$^2$ in plane resolution).

Data were viewed and analyzed using AFNI and related plug-in software-http://afni.nimh.nih.gov/ [Cox, 1996]. Brain volumes were motion corrected for global head translations and rotations by registering them to the 20th brain volume of the functional scan closest to the anatomical scan. BOLD signals were converted to percent signal changes on a voxel-by-voxel basis relative to responses to the silent events within each scan. For each participant, the functional scans (seven separate runs) were concatenated into a single time series. We then performed multiple linear regression analyses based on the button responses modeling whole-brain BOLD signal responses to the sound stimuli relative to the baseline silent events. With the clustered acquisition design, the BOLD response to each sound stimulus could be treated as an independent event. In particular, brain responses to stimulus events could be censored from the model in accordance with each participant's button responses. Because participants were instructed to listen to the entire sound sample and respond as accurately as possible (3AFC) as to their sense of recognition, rather than as fast as possible, an analysis of reaction times was not reliable. For the main analysis of each individual dataset, we included only those sound pairs where the forward-played version was judged to be recognizable (RF, Recognized Forward; button no. 1) and the corresponding backward-played version was judged as not being recognizable (NB, Not recognized Backward; button no. 3).

For this study, several sound stimulus events were additionally censored *post hoc* from all analyses for all individuals. This included censoring nine of the sound stimuli that contained vocalization content, which were excluded to avoid confounding network activation associated with pathways that may be specialized for processing vocalizations [Belin et al., 2004; Lewis et al., 2009]. Additionally, we subsequently excluded sounds that were not directly associated with a human agent instigating the action. To assess human agency, five age-matched sighted participants not included in the fMRI scanning rated all sounds on a Likert scale (1–5) indicating whether the heard sound (without reference to any verbal or written descriptions) evoked the sense that a human was directly involved in the sound production (1 = human, 3 = not sure, 5 = not associated with human action). Sounds with an average rating between 1 and 2.5 were analyzed separately as human-produced action sounds, resulting in retaining 61 of the 108 sound stimuli (Appendix).

Using AFNI, individual anatomical and functional brain maps were transformed into the standardized Talairach coordinate space [Talairach and Tournoux, 1988]. Functional data (multiple regression coefficients) were spatially low-pass filtered (6-mm box filter), then merged by combining coefficient values for each interpolated voxel across all subjects. A voxel-wise two sample $t$-test was performed using the RF versus NB regression coefficients to identify regions showing significant differences between the EB and SC groups. The results from this comparison were re-

stricted to reveal only those regions showing positive differences in RF versus NB comparisons (i.e., where RF sounds led to greater positive BOLD signal changes in at least one of the two groups of listeners). This approach excluded regions differentially activated solely due to differential negative BOLD signals, wherein the unrecognized backward-played sounds evoked greater magnitude of activation relative to recognized sounds for both groups. Corrections for multiple comparisons were based on a Monte Carlo simulation approach implemented by AFNI-related programs AlphaSim and 3dFWHM. A combination of individual voxel probability threshold ($t$-test, $P < 0.02$ or $P < 0.05$; see Results) and the cluster size threshold (12 or 20 voxel minimum, respectively), based on an estimated 2.8 mm$^3$ full-width half-max spatial blurring (before low-pass spatial filtering) present within individual datasets, yielded the equivalent of a whole-brain corrected significance level of $\alpha < 0.05$.

Data were then projected onto the PALS atlas brain database using Caret software-http://brainmap.wustl.edu [Van Essen, 2005; Van Essen et al., 2001]. Surface-registered visual area boundaries—e.g., V1 and V2 [Hadjikhani et al., 1998]—from the PALS database were superimposed onto the cortical surface models, as were the reported coordinate and approximated volumetric locations of the left and right parahippocampal place area (PPA) [Epstein and Kanwisher, 1998; Gron et al., 2000]. Portions of these data can be viewed at http://sumsdb.wustl.edu/sums/directory.do?id = 6694031&dir_name = LEWIS_HBM10, which is part of a database of surface-related data from other brain mapping studies.
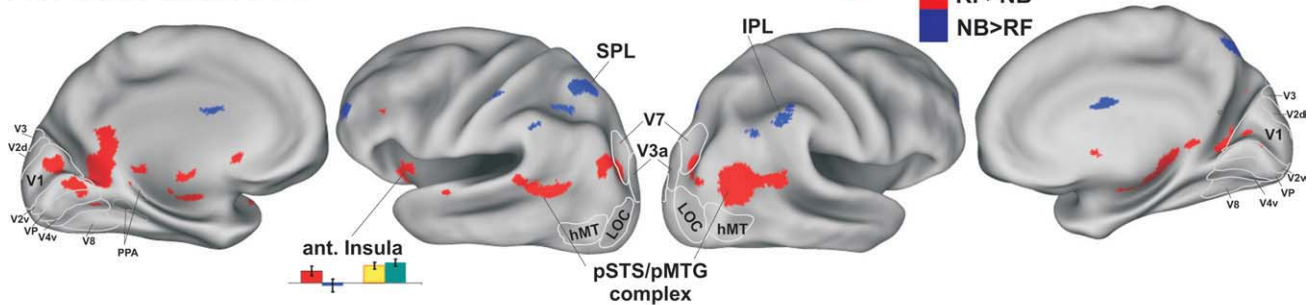
## RESULTS

To reveal brain regions preferentially involved in the process of recognizing the human-produced action sounds, for each individual, we effectively subtracted the activation response to the unrecognized, backward-played sounds from that for the corresponding recognized, forward-played human action sounds (refer to Methods). There was no statistical difference in the distribution of numbers of retained sound pairs between the two groups (see Appendix inset; 30.6% for EB, 35.9% for SC; two sample, two-tailed $t$-test, $T_{(22)} = 1.48$; $P > 0.15$), indicating that the EB and SC groups were comparable overall in their ability to recognize the particular everyday sound stimuli retained for analyses. The resulting group-averaged patterns of cortical activation to recognized forward-played sounds relative to unrecognized backward-played sounds are illustrated on the PALS cortical surface model for the SC group (Fig. 1A, yellow vs. dark green, $t$-test; $T_{(14)} = 2.65$, $P < 0.02$, 12 voxel minimum cluster size correcting to $\alpha < 0.05$) and the EB group (Fig. 1B, red vs. blue; $T_{(10)} = 2.82$, $P < 0.02$, $\alpha < 0.05$). Although both the forward- and backward-played sounds strongly activated cortex throughout most of auditory cortex proper—including primary auditory cortices residing along
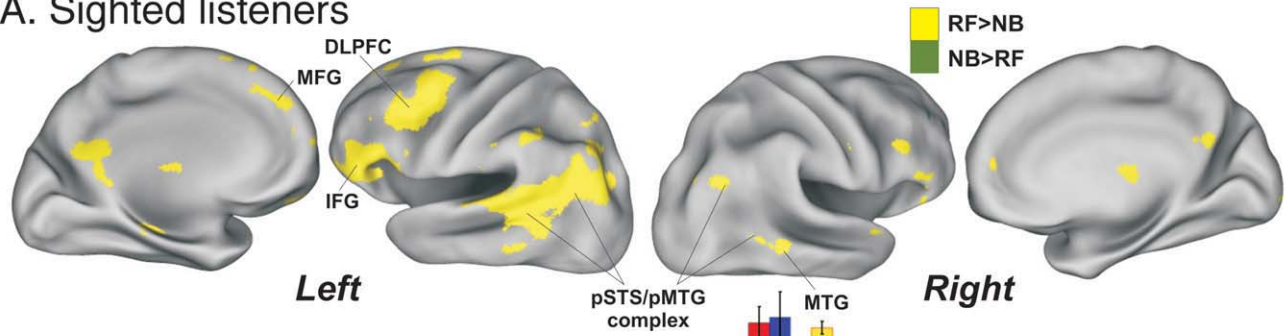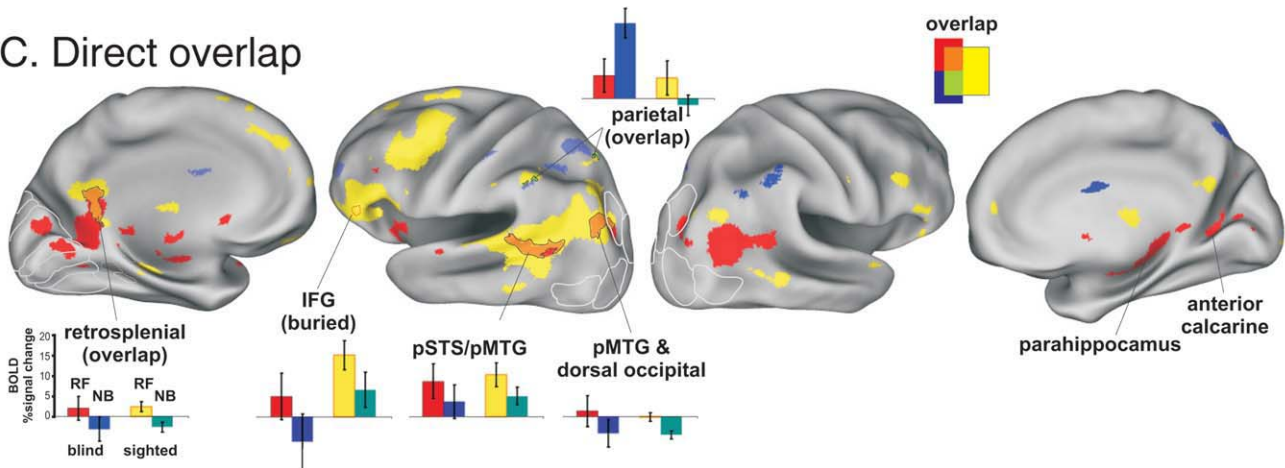
**Figure 1.**

Comparison of cortical networks for sighted versus early blind listeners associated with recognition of human action sounds. All results are illustrated on 3D renderings of the PALS cortical surface model, and all colored foci are statistically significant at α < 0.05, whole brain corrected. **A**: Group-averaged activation results from age-matched sighted participants (*n* = 14) when hearing and recognizing forward-played human action sounds (RF, yellow) in contrast to not recognized backward-played versions of those sounds (NB, green). **B**: Group-averaged activation in the early blind (*n* = 10) for recognized, forward-played (RF, red) versus unrecognized backward-played (RB, blue) sound stimuli. **C**: Regions showing direct overlap with the same (orange) or opposite (green) differential activation pattern. All histograms illustrate responses to recognized forward-played (RF) sounds and the corresponding backward-played sounds not recognized (NB) relative to silent events for both groups. Outlines of visual areas (V1, V2, V3, hMT, etc., white outlines) are from the PALS atlas database. CaS, calcarine sulcus; CeS, central sulcus; CoS, collateral sulcus; IFG, inferior frontal gyrus; IPS, intraparietal sulcus; STS, superior temporal sulcus. The IFG focus did not directly intersect the cortical surface model, and thus an approximate location is indicated by dotted outline. Refer to color inset for color codes, and the text for other details. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
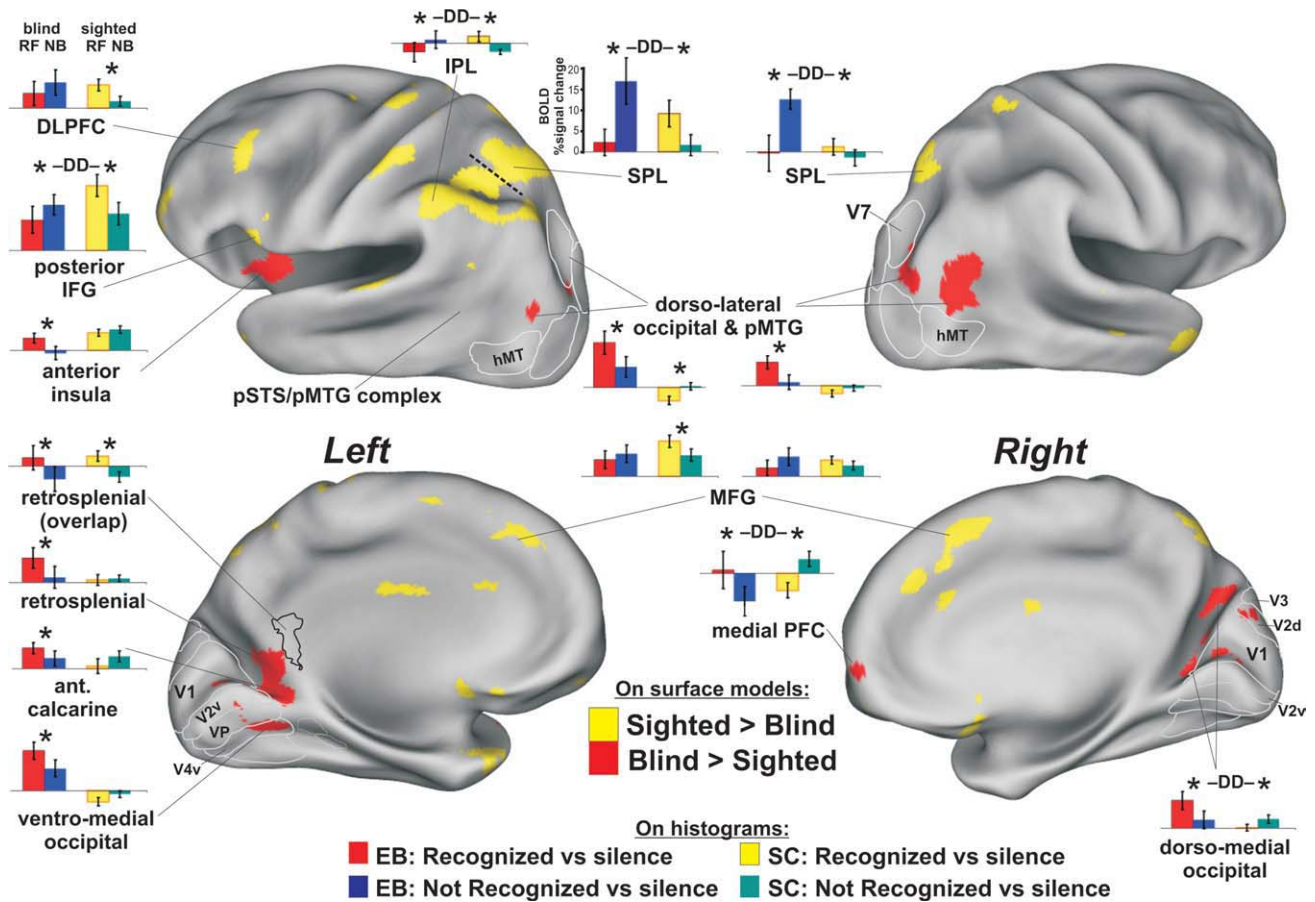
**Figure 2.**

Significant differences between the sighted control group (yellow on surface models) and early blind (red) group ($\alpha < 0.05$, corrected). Foci emerging solely due to differences of negative BOLD signals were excluded. Asterisks denote significant two-tailed *t*-test differences at $P < 0.05$. The DD icons designate ROIs showing a significant double dissociation of response profiles between groups. Refer to Figure 1 and text for other details. EB = early blind group; SC = sighted control group. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Heschl's gyrus and the superior temporal plane [Formisano et al., 2003; Lewis et al., 2009; Rademacher et al., 2001; Talavage et al., 2004]—no significant differential activation was observed in these regions. Consistent with our earlier report using this basic paradigm [Lewis et al., 2004], this was most likely due to the close match in physical attributes between the forward- versus backward-played sounds (refer to Methods). Rather, the main difference between the two sets of sound stimulus pairs was that only the forward-played sounds were judged as being recognizable. Thus, the cortical activations revealed after the different sets of subtractions (Figs. 1 and 2, colored brain regions; also see Tables I and II) should largely reflect high-level perceptual or conceptual processing associated with the participants' judgments of having recognized the acoustic events. Note that neither the dorso-lateral prefrontal cortex (DLPFC) nor medial frontal gyrus (MFG) regions (Fig. 1A, yellow) were differentially activated in our earlier study (ibid). It remains unclear whether this was due to use of a lower field strength MRI scanner (1.5T vs. 3T), lower spatial resolution, the particular sounds retained for analysis, a younger group of sighted listeners (mostly 20–30 years of age), or some combination therein. Nonetheless, in this study, a similar distribution of sounds were retained for both the sighted and blind group, and the two groups were carefully matched in ages to minimize that as a source of variability, which is known to impact fMRI activation patterns in both SC and EB listeners [Roder and Rosler, 2003].

Regions showing direct overlap between the SC and EB groups that were preferentially activated by the forward-played sounds judged to have been recognized (Fig. 1C; intermediate color orange) included the anterior and posterior portions of the left pSTS/pMTG complex, left inferior frontal gyrus (IFG, dotted outline), and left retrosplenial cortex. Both groups also revealed regions that showed greater responses to the unrecognized backward-

**TABLE I. Talairach coordinates of several sound recoginition foci, corresponding to Figure 1**

| | Anatomical location | Talairach coordinates | | | Volume (mm³) |
|---|---|---|---|---|---|
| | | $x$ | $y$ | $z$ | |
| **Right hemisphere** | | | | | |
| Early Blind | Medial occipital | 14 | −76 | 15 | 2916 |
| | pMTG/dorso-lateral occipital | 44 | −65 | 17 | 4937 |
| Sighted Controls | pMTG, dorsal occipital | 61 | −56 | 5 | 4438 |
| **left hemisphere** | | | | | |
| Early Blind | Ventral-Medial Occ | −13 | −69 | −3 | 4495 |
| | Dorsal Occipital foci | −36 | −78 | 23 | 3331 |
| | Anterior Calcarine | −11 | −53 | 4 | 3261 |
| Both SC and EB | pMTG/pSTS | −53 | −53 | 11 | 1027 |
| | pMTG, dorsal occipital | −42 | −67 | 20 | 2815 |
| | IFG (overlap) | −54 | 35 | 7 | 63 |
| | IPL | −36 | −75 | 26 | 1777 |
| | Restrosplenial | −7 | −50 | 14 | 464 |
| Sighted Controls | IFG | −47 | 23 | 11 | 897 |
| | DLPFC | −45 | 11 | 31 | 1530 |
| | pSTS/pMTG | −49 | −55 | 20 | 26885 |

Refer to text acronyms.

played sounds relative to the recognized forward-played sounds, each relative to silent events. For the EB group (Fig. 1B, blue), this included activation along the bilateral inferior parietal lobule (IPL) and left superior parietal lobule (SPL) regions. In these parietal regions, the response profiles of the SC group were opposite to those of the EB group (e.g., Fig. 1C, light green ROIs and histogram). In particular, the SC group revealed significant positive, differential activation to recognized human action sounds and little or no response to the corresponding unrecognized backward-played sounds (yellow and dark green in histogram), whereas the EB group (red and blue) showed significant positive differential activation to backward-played sounds deemed as being unrecognizable, yet little or no activation to the forward-played sounds judged as having been recognized.

Significant differences in the global activation patterns between the EB and SC groups were directly assessed in a second analysis using a voxel-wise two sample $t$-test between groups (Fig. 2, yellow vs. red, $T_{(10,14)} = 2.82$, $P < 0.05$ with 20 voxel cluster minimum corrected to $\alpha < 0.05$). At a gross level, the SC group preferentially activated

**TABLE II. Talairach coordinates of activation foci significantly different between the sighted and blind groups, corresponding to Figure 2**

| | Anatomical location | Talairach coordinates | | | Volume (mm³) |
|---|---|---|---|---|---|
| | | $x$ | $y$ | $z$ | |
| **Right hemisphere** | | | | | |
| Early Blind | pMTG, dorso-lateral occipital | 37 | −72 | 19 | 4,050 |
| Sighted Controls | Medial frontal gyrus | 6 | 22 | 42 | 2,645 |
| | SPL | 21 | −72 | 53 | 2,037 |
| | Medial PFC | 9 | 63 | 2 | 342 |
| **left hemisphere** | | | | | |
| Early Blind | Dorso-lateral occipital | −34 | −81 | 13 | 1,438 |
| | Anterior Insula | −38 | 5 | −3 | 979 |
| | Anterior calcarine | −13 | −57 | 8 | 1,693 |
| | Dorso-medial occipital | −8 | −75 | 4 | 287 |
| | Ventro-medial occipital | −21 | −66 | −3 | 4,233 |
| Sighted Controls | SPL | −34 | −62 | 58 | 12,794 |
| | IPL | −43 | −56 | 43 | 4,345 |
| | DLPFC | −48 | 22 | 29 | 1,229 |
| | Posterior IFG | −52 | 12 | 4 | 1,166 |
| | Medial frontal gyrus | −4 | 27 | 41 | 2,450 |

dorsal cortical regions, including several lateral and medial portions of parietal and frontal cortex. In contrast, the EB group preferentially activated dorso-lateral occipital regions located immediately posterior to the bilateral pSTS/pMTG complexes, the left anterior insula and portions of the nearby left basal ganglia (not shown) together with bilateral medial occipital and anterior calcarine cortices (Fig. 2, red). Thus, for the EB group, there was an overall greater expanse of activation in and around the pMTG regions, especially in the right hemisphere, in response to the judgment of recognition of the human action sounds.

In medial occipital regions, the preferential activation by the EB group included cortices that would otherwise have developed to become retinotopically organized visual areas, including estimated locations (white outlines) of V1, V2d, V2v, VP, and V4v, plus the dorso-lateral visual area V7 [Grill-Spector et al., 1998; Hadjikhani et al., 1998; Shipp et al., 1995]. The activated medial occipital regions were largely restricted to cortex that would typically represent peripheral visual field locations [Brefczynski and DeYoe, 1999; Engel et al., 1994; Sereno et al., 1995]. For the EB group, the left hemisphere anterior calcarine and ventro-medial occipital foci were largely contiguous (in volumetric space) with the retrosplenial focus, though there were some subtle differences in their respective response profiles (Fig. 2, histograms). For instance, all occipital ROIs preferentially activated by the EB group showed significant responses to both forward- and backward-played sounds relative to silent events. However, for the SC group, none of the occipital ROIs showed significant activation to recognizable forward-played sounds.

Interestingly, when the SC group could not readily recognize the action sounds, there was relatively greater activation along some of the midline cortical regions preferentially activated during successful sound recognition by the EB group (Fig. 2, green vs. yellow histograms for network of red cortical regions). This was statistically significant for the right dorso-medial occipital cortex ($t_{(26)}$ = 1.74, $P < 0.05$) and right medial prefrontal cortex ($t_{(26)}$ = 3.2, $P < 0.004$) and showed at least a trend in left ventromedial occipital and anterior calcarine cortex. Thus, in comparison with histograms in fronto-parietal regions (i.e., IPL, SPL, and IFG), there was evidence of a double dissociation (DD labels over histograms) of response profiles between these two groups of listeners.

## DISCUSSION

There were two major findings from this study. First, answering our original question regarding audio–visual functions of the bilateral pSTS/pMTG complexes, both the sighted control (SC) and early blind (EB) groups did significantly activate substantial portions of these regions in response to hearing and recognizing human action sounds in contrast to unrecognized, acoustically well-matched back-

ward-played sounds as the control condition. This demonstrated that even in the complete absence of visual experience, the left and right postero-lateral temporal regions can develop to subserve functions related to biological and/or complex motion processing that are conveyed through the auditory system. Second, the SC and EB groups additionally revealed distinct differences in global network activation in response to the judgment of recognition of human action sounds: the SC group preferentially activated left-lateralized IPL and frontal regions, whereas the EB preferentially or uniquely activated lateral and medial occipital regions and left anterior insular cortex. These results were suggestive of different strategies for recognition or memory recall of human action sounds. Below, we address possible functional roles of cortical systems commonly activated by both groups (i.e., Fig. 1C, orange regions), followed by discussion of two potentially distinct cortical mechanisms for representing acoustic knowledge of human action sounds (i.e., Fig. 2, yellow vs. red networks).

## The pSTS/pMTG Complexes and Action Knowledge Representations

The original rationale for this study was to characterize the functional roles of the pSTS/pMTG regions. From our earlier study, these lateral temporal complexes appeared consistent with having functions related to visual imagery that may have been directly or indirectly evoked by the real-world acoustic events [Lewis et al., 2004]. As the EB participants in this study have never had significant visual object or visual motion experiences, the activation of the left pSTS/pMTG regions by both sighted and blind participants indicated that the human action sounds were not simply evoking or representing correlates of "visual" imagery per se. Thus, if the pSTS/pMTG complexes are involved in some aspect of mental imagery conjured up by the sounds, then this would not be consistent with solely vision-based Depictive View theories, which entail picture or picture-like representations in cortex [Slotnick et al., 2005]. Rather, the present findings would be compatible with either more abstract representations of action dynamics, such as with Symbol-Structure theories of reasoning and thought [Pylyshyn, 2002], or possibly a modified Depictive View theory that is more multisensory-inclusive. Mental imagery notwithstanding, we regard the functions of the pSTS/pMTG complexes to have prominent perceptual roles in transforming spatially and temporally dynamic features of familiar, human-produced auditory or visual action information together into a common neural code [for review see Lewis, 2010].

Two extremes in theories for the general organization of knowledge representations in human cortex include domain-specific and sensory-motor property-based models [Martin, 2007]. In general, domain-specificity theories posit that some cortical regions are genetically predisposed (or "hard-wired") to perform certain types of processing

operations independent of an individual's sensory experience or capabilities [Caramazza and Shelton, 1998; Mahon and Caramazza, 2005]. The similar response profiles along the left pSTS/pMTG complexes by both the SC and EB groups were consistent with these regions performing "metamodal" operations [Pascual-Leone and Hamilton, 2001] that may reflect domain-specific hubs. Thus, biological motion processing may be added to a growing list of cortical operations common to sighted and blind listeners, including object shape processing [Amedi et al., 2007], conceptual level and linguistic functions [Burton and McLaren, 2006; Burton et al., 2002a,b; Roder et al., 2002], and processing of category-specific object knowledge [Mahon et al., 2009]. Thus, large expanses of cortical regions or networks may be operating intrinsically, such that visual or acoustic information, if present, modulates rather than determines the functional operations [Burton et al., 2004].

Alternatively, or additionally, the functions of the pSTS/pMTG complexes may develop in part due to the greater familiarity we have with human action sounds [Lewis et al., in press] and thus be driven by sensory-motor property-based mechanisms [Kiefer et al., 2007; Lissauer, 1890/1988; Martin, 2007; Warrington and Shallice, 1984]. In particular, there is considerable survival value in learning to associate characteristic sounds with the behaviorally relevant motor actions that produce them (and the associated visual motion characteristics), notably including the actions of one's caretakers and one's self experienced throughout early stages of development. Regardless of which model predominates, the bilateral pSTS/pMTG regions appear to develop as hubs for human biological motion processing independent of sensory input, thus serving to shape the cortical organization of representations related to action knowledge, and perhaps even more abstract-level processing such as linguistic representations.

Cortex overlapping and located immediately posterior to the bilateral pSTS/pMTG complexes have been reported to be activated during spoken language processing in blind individuals, especially with more difficult syntax and semantic content [Roder et al., 2002]. Additionally, in sighted participants, the left pSTS/pMTG regions are activated by words and sentences depicting actions, notably including human conspecific actions [Aziz-Zadeh et al., 2006; Kellenbach et al., 2003; Kiefer et al., 2008; Tettamanti et al., 2005]. Consistent with a grounded cognition framework, linguistic and/or conceptual-level representations of action events may largely be housed in the same networks involved in processing the associated auditory, visual, or sensory-motor events [Barsalou, 2008; Barsalou et al., 2003; Kiefer et al., 2008]. As such, action events, whether viewed, heard, audio–visual associated, or reactivated in response to imagery or verbal depictions of those action events, may be probabilistically encoded or "grounded" in regions such as the pSTS/pMTG complexes. This may ultimately serve to mediate both perceptual and conceptual action knowledge representations.

## Different Networks for Human Action Sound Recognition in Sighted Versus Blind Listeners

Although the sighted and early blind individuals did not report any obvious differences in "how" they were able to recognize the action sounds we presented, the distinct differences in their respective network activations (Fig. 2, yellow vs. red) are suggestive of different retrieval or sound recognition strategies. We propose that the two networks revealed by the SC group (yellow) and the EB group (red) reveal two distinct cortical network systems that can be utilized to subserve memory retrieval related to recognition of acoustic events. Four cortical subnetworks that appear to encompass these global recognition networks include left-lateralized fronto-parietal cortices, bilateral SPL regions, the left anterior insula, and posterior midline structures, which are addressed in turn below.

### Left IPL and IFG

The left-lateralized IPL and IFG cortices preferentially activated by the SC group appeared to overlap with reported mirror-neuron systems [Gallese and Goldman, 1998; Rizzolatti and Craighero, 2004; Rizzolatti et al., 1996]. These systems are involved in both observation and execution of meaningful motor actions and represent a potential mechanism for how an observer may attain a sense of meaning behind viewed and/or heard actions. A recent neuroimaging study also comparing sighted with early blind listeners indicated that auditory mirror systems (left IPL and IFG) develop in the absence of visual experience, such that familiar hand-executed action sounds can evoke motor action schemas not learned or experienced through the visual modality [Ricciardi et al., 2009]. Consistent with their study, the judgment of recognition of human action sounds by sighted listeners in this study seemed likely to be associated with reasoning about how the sound was produced in terms of one's own motor repertoires and audio-motor (and perhaps audio–visual-motor) associations. Surprisingly, however, the EB group in this study showed the opposite response profile results in left IPL and IFG regions (and bilateral SPL), revealing activation that was significantly preferential for the unrecognized backward-played sounds, and a trend toward this response profile in the left DLPFC and bilateral MFG regions (Fig. 2, histograms).

Although the results of this study do not refute the seemingly contradictory results by Ricciardi et al., [2009], they do indicate that fronto-parietal networks (presumed mirror-neuron systems) may be evoked only under particular task conditions to help provide a sense of meaning for certain action sounds. There were several differences between the two studies that could account for these discrepant results. First, this study incorporated a wider range of human action sounds, including lower body, whole body, and mouth, in addition to hand-executed action sounds. Second, none of our sounds were repeated

during the scanning session, thereby reducing possible repetition suppression or enhancement effects due to learning through familiarization [Grill-Spector et al., 2006]. Third, we used unrecognized backward-played sound stimuli as a control condition to maximize the effects associated with perceived recognition, as opposed to using motor familiarity ratings. Lastly, and perhaps most importantly, was that our paradigm did not include pantomiming or any motor-related tasks. We suspect that the inclusion of motor-related tasks during or close in time to a listening task may prime one's motor system, thereby driving increased recruitment and activation of left-lateralized fronto-parietal networks. Thus, the recognition requirements of the present paradigm may have permitted blind listeners to use alternative or "preferred" recall strategies not related to motor inferencing.

### Bilateral SPL

The left and right SPL regions represent another differentially activated system between the sighted and blind groups (Fig. 2, yellow). In the left hemisphere, the SPL and IPL activation foci were contiguous, but were separated (dashed line, based on the right SPL location) for purposes of generating distinct ROI histograms. Not only were the bilateral SPL regions preferentially activated by recognized human action sounds in the SC group (yellow vs. green), but more strikingly the EB group showed the most robust activation in response to the backward-played unrecognized sounds (blue vs. red histograms). These SPL regions may have had a role in memory retrieval. For instance, parietal cortex is reported to contribute to episodic memory functions [Wagner et al., 2005], including attention directed at internal mnemonic representations, to the eventual decision of recognition success [Yonelinas et al., 2005], and/or to working memory buffers for storing information involved in the recollection process [Baddeley, 1998]. A more recent account of parietal function includes an attention to memory (AtoM) theory, suggesting that these regions may be evoked when pre- and postretrieval processing is needed to produce a memory decision for the task at hand [Ciaramelli et al., 2008, 2010]. In particular, they hypothesized that SPL regions allocate top–down attentional resources for memory retrieval. This is further supported by a study of patients with bilateral parietal lobe lesions, who in general experienced reduced confidence in episodic recollection abilities [Simons et al., 2010]. Why the bilateral SPL system may have been evoked more by the EB group during failed recognition is considered after first addressing the left anterior insula and medial occipital networks, which were preferentially activated during perceived recognition success.

### Left anterior insula

In contrast to the SC group, perceived sound recognition by the EB group preferentially evoked activity in the left anterior insula, which is reported to be associated with the "feeling of knowing" before recall in the context of meta-memory monitoring [Kikyo et al., 2002]. More generally, this region is proposed to play a crucial role in interoception, and the housing of internal models or meta-representations of one's own behaviors as well as modeling the behaviors of others [Augustine, 1996; Craig, 2009; Mutschler et al., 2007]. Thus, differential activation of the left anterior insula may have been associated with episodic memory for situational contexts, including representations of "self" and processing related to social cognition.

### Medial occipital cortices

In response to recognized sounds, the EB group uniquely revealed a nearly contiguous swath of activity that extended from the commonly activated left retrosplenial cortex to the bilateral anterior calcarine and medial occipital regions. One intriguing possibility is that this apparent "expansion" of retrosplenial activation (cf. Fig. 1C orange and red) may reflect the same basic type of high-level processing. The retrosplenial cortex together with surrounding posterior cingulate and anterior calcarine regions have been implicated in a variety of critical functions related to episodic memory [Binder et al., 1999; Kim et al., 2007; Maddock, 1999; Valenstein et al., 1987]. This includes parametric sensitivity to the sense of memory confidence and retrieval success [Heun et al., 2006; Kim and Cabeza, 2009; Mendelsohn et al., 2010; Moritz et al., 2006], the sense of perceptual familiarity with complex visual scenes [Epstein et al., 2007; Montaldi et al., 2006; Walther et al., 2009], and with aspects of self appraisal and autobiographical knowledge [Burianova et al., 2010; Johnson et al., 2002; Ries et al., 2006, 2007; Schmitz et al., 2006]. Together with the medial PFC (e.g., Fig. 2, red) and other frontal regions, retrosplenial activation is frequently implicated in social cognitive functions, theory of mind, knowledge of people, and situational contexts involving people [Bedny et al., 2009; Maddock et al., 2001; Simmons et al., 2009; Walter et al., 2004; Wakusawa et al., 2007, 2009]. While the functions of retrosplenial and surrounding cortex remain to be more precisely understood, the present results suggest that blind listeners are more apt to encode representations of human action sounds using this sub-network, which relate less to motor intention inferencing and more to situational contexts.

Regarding the reorganization of "visual cortex" in the blind, the EB group uniquely activated several midline occipital regions that overlapped with what would otherwise have developed to become retinotopically organized visual areas [Kay et al., 2008; Tootell et al., 1998]. One proposed mechanism for such plasticity is that in the absence of visual input during development, fronto-parietal networks that feed into occipital cortex modulate neural responses as a function of attention in a general manner, allowing for more interactions between salient acoustic and/or tactile-motor events [Stevens et al., 2007; Weaver

and Stevens, 2007]. Another mechanism that may be germane to cortical reorganization of medial occipital cortices entails a "Reverse Hierarchy" model. In the context of the visual system, this model stipulates that the development of a neural system for object recognition and visual perception learning is a top-down guided process [Ahissar et al., 2009; Hochstein and Ahissar, 2002]. The retrosplenial and anterior calcarine regions may serve as such high-level supramodal or amodal processing regions. In particular, the anterior calcarine sulci were shown to have marked functional connectivity with primary auditory cortices in humans, and was hypothesized to mediate cross-modal linkages between representations of peripheral field locations in primary visual areas and auditory cortex [Eckert et al., 2008]. Similarly, in the macaque monkey, the anterior calcarine region is reported to represent a multisensory convergence site [Rockland and Ojima, 2003]. Thus, the anterior calcarine may generally function to represent aspects of high-level scene interpretation, independent of sensory modality. In sighted individuals, linkages to both early visual and early auditory cortices may permit the encoding of lower-level units of detailed information for the respective sensory modality input. However, in the absence of visual experience, the anterior calcarine may in part be driving the apparent cortical reorganization of occipital regions to subserve functions related to acoustic scene action knowledge, which would thus include high-level representations of episodic memories and/or situational relationships learned through auditory and audio-tactile-motor experience.

### Network models of sound recognition

Regardless of whether domain-specific or sensory-motor property-based encoding mechanisms prevail, the formation of global cortical network differences between the SC and EB groups for human action sound recognition described above may be regarded in the broader context of pattern recognition models (e.g., Hopfield networks). In particular, such neural networks may activate or settle in one of potentially many different states that would provide the listener with different forms (or degrees) of perceived recognition success [Hopfield and Brody, 2000; Hopfield and Tank, 1985]. Thus, acoustic knowledge representations could manifest as local minima or basins within large-scale cortical networks that may develop to be differently weighted depending on whether or not the individual has had visual experiences to associate with acoustic events. Sighted listeners will have had life-long visual exposure to sound-producing actions of other humans (conspecifics), thereby strengthening audio–visual-motor associations (via Hebbian-like mechanisms). Consequently, sighted individuals may rely more heavily on motor-inferencing strategies to efficiently query probabilistic matches between the heard action sounds and their own repertoires of motor actions or schemas (and visual motion dynamics) that produce characteristic sounds. This would be consistent with the activation of left-lateralized fronto-parietal

mirror-neuron systems. In contrast, the early blind participants may have relied less on motor inferencing and more on probabilistic matches to other situational context memories to provide a sense of recognition of the action events-involving representations housed in left anterior insular cortex and bilateral medial occipital regions. However, when the blind listeners were presented with backward-played sounds that they did not readily recognize or match to previously learned situational contexts, they may have then resorted to utilizing bilateral SPL attentional systems either to help query fronto-parietal mirror-neuron systems (attempting a motor inferencing strategy) or to search and monitor for other operations needed to inform recollective decisions. While a double-dissociation in activation profiles existed for these two apparently distinct recognition or memory retrieval systems (between the sighted and blind listeners), further study examining the temporal dynamics of these processing pathways will be needed to validate this interpretation.

In sum, the present results demonstrated that human-produced biological action sounds commonly activate large portions of the bilateral pSTS/pMTG complexes and left retrosplenial cortex in both sighted and blind listeners, the latter of whom have never had visual experiences to associate with those sound-source actions. Thus, the pSTS/pMTG regions appear to represent domain-specific hubs for processing biological or complex dynamic motion attributes that can develop independent of visual experience. In comparing sighted versus blind listeners, striking differences in global network activation patterns to recognized versus unrecognized human action sounds were also observed. Sighted listeners relied more heavily on left-lateralized fronto-parietal networks associated with mirror-neuron systems, consistent with a motor inferencing strategy for attaining a sense of recognition of human action sounds. In contrast, blind listeners more heavily utilized the left anterior insula and medial occipital regions, consistent with memory retrieval related to scene representations and situational relationships. In the blind group, the unique recruitment of lateral occipital regions that were juxtaposed to the pMTG, and medial occipital regions that were contiguous with retrosplenial cortex, was suggestive of expansions of cortical function related to action processing and other knowledge representations of human-produced action sounds. Thus, the absence of visual input appears to not only reweight which global cortical network mechanisms are preferentially used for acoustic knowledge encoding but also permits or guides the corresponding functional expansions and reorganization (cortical plasticity) of occipital cortices.

# REFERENCES

Ahissar M, Nahum M, Nelken I, Hochstein S (2009): Reverse hierarchies and sensory learning. Philos Trans R Soc Lond B Biol Sci 364:285–299.

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, 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.

Augustine JR (1996): Circuitry and functional aspects of the insular lobe in primates including humans. Brain Res Brain Res Rev 22:229–244.

Aziz-Zadeh L, Wilson SM, Rizzolatti G, Iacoboni M (2006): Congruent embodied representations for visually presented actions and linguistic phrases describing actions. Curr Biol 16:1818–1823.

Baddeley A (1998): Recent developments in working memory. Curr Opin Neurobiol 8:234–238.

Barsalou LW (2008): Grounded cognition. Annu Rev Psychol 59:617–645.

Barsalou LW, Kyle Simmons W, Barbey AK, Wilson CD (2003): Grounding conceptual knowledge in modality-specific systems. Trends Cogn Sci 7:84–91.

Bavelier D, Tomann A, Hutton C, Mitchell T, Corina D, Liu G, Neville H (2000): Visual attention to the periphery is enhanced in congenitally deaf individuals. J Neurosci 20:RC93.

Beauchamp M, Lee K, Haxby J, Martin A (2002): Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34:149–159.

Beauchamp MS, Lee KM, Argall BD, Martin A (2004a): Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809–823.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004b): Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190–1192.

Bedny M, Pascual-Leone A, Saxe RR (2009): Growing up blind does not change the neural bases of Theory of Mind. Proc Natl Acad Sci USA 106:11312–11317.

Belin P, Fecteau S, Bedard C (2004): Thinking the voice: Neural correlates of voice perception. Trends Cogn Sci 8:129–135.

Bidet-Caulet A, Voisin J, Bertrand O, Fonlupt P (2005): Listening to a walking human activates the temporal biological motion area. Neuroimage 28:132–139.

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW (1999): Conceptual processing during the conscious resting state: A functional MRI study. J Cogn Neurosci 11:80–95.

Brefczynski JA, DeYoe EA (1999): A physiological correlate of the 'spotlight' of visual attention. Nat Neurosci 2:370–374.

Burianova H, McIntosh AR, Grady CL (2010): A common functional brain network for autobiographical, episodic, and semantic memory retrieval. Neuroimage 49:865–874.

Burton H (2003): Visual cortex activity in early and late blind people. J Neurosci 23:4005–4011.

Burton H, McLaren DG (2006): Visual cortex activation in late-onset, Braille naive blind individuals: An fMRI study during semantic and phonological tasks with heard words. Neurosci Lett 392:38–42.

Burton H, Snyder AZ, Raichle ME (2004): Default brain functionality in blind people. Proc Natl Acad Sci USA 101:15500–15505.

Burton H, Snyder AZ, Diamond JB, Raichle ME (2002a): Adaptive changes in early and late blind: A FMRI study of verb generation to heard nouns. J Neurophysiol 88:3359–3371.

Burton H, Snyder AZ, Conturo TE, Akbudak E, Ollinger JM, Raichle ME (2002b): Adaptive changes in early and late blind: A fMRI study of Braille reading. J Neurophysiol 87:589–607.

Calvert GA, Brammer MJ (1999): FMRI evidence of a multimodal response in human superior temporal sulcus. Neuroimage 9:S1038.

Calvert GA, Campbell R (2003): Reading speech from still and moving faces: The neural substrates of visible speech. J Cogn Neurosci 15:57–70.

Calvert GA, Campbell R, Brammer MJ (2000): Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10:649–657.

Calvert GA, Brammer MJ, Bullmore ET, Campbell R, Iversen SD, David AS (1999): Response amplification in sensory-specific cortices during crossmodal binding. NeuroReport 10:2619–2623.

Caramazza A, Shelton JR (1998): Domain-specific knowledge systems in the brain the animate-inanimate distinction. J Cogn Neurosci 10:1–34.

Ciaramelli E, Grady CL, Moscovitch M (2008): Top-down and bottom-up attention to memory: A hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. Neuropsychologia 46:1828–1851.

Ciaramelli E, Grady C, Levine B, Ween J, Moscovitch M (2010): Top-down and bottom-up attention to memory are dissociated in posterior parietal cortex: Neuroimaging and neuropsychological evidence. J Neurosci 30:4943–4956.

Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Craig AD (2009): How do you feel--now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Eckert MA, Kamdar NV, Chang CE, Beckmann CF, Greicius MD, Menon V (2008): A cross-modal system linking primary auditory and visual cortices: Evidence from intrinsic fMRI connectivity analysis. Hum Brain Mapp 29:848–857.

Edmister WB, Talavage TM, Ledden PJ, Weisskoff RM (1999): Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 7:89–97.

Engel LR, Frum C, Puce A, Walker NA, Lewis JW (2009): Different categories of living and non-living sound-sources activate distinct cortical networks. Neuroimage 47:1778–1791.

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky E, Shadlen MN (1994): fMRI of human visual cortex. Nature 369:525.

Epstein R, Kanwisher N (1998): A cortical representation of the local visual environment. Nature 392:598–601.

Epstein RA, Higgins JS, Jablonski K, Feiler AM (2007): Visual scene processing in familiar and unfamiliar environments. J Neurophysiol 97:3670–3683.

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R (2003): Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40:859–869.

Gallese V, Goldman A (1998): Mirror neurons and the simulation theory of mind-reading. Trends Cogn Sci 2:493–501.

Gazzola V, Aziz-Zadeh L, Keysers C (2006): Empathy and the somatotopic auditory mirror system in humans. Curr Biol 16:1824–1829.

Glover GH, Law CS (2001): Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magn Reson Med 46:515–522.

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.

Grill-Spector K, Henson R, Martin A (2006): Repetition and the brain: Neural models of stimulus-specific effects. Trends Cogn Sci 10:14–23.

Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzchak Y, Malach R (1998): A sequence of object-processing stages revealed by fMRI in the human occipital lobe. Hum Brain Mapp 6:316–328.

Gron G, Wunderlich AP, Spitzer M, Tomczak R, Riepe MW (2000): Brain activation during human navigation: Gender-different neural networks as substrate of performance. Nat Neurosci 3:404–408.

Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R (2000): Brain areas involved in perception of biological motion. J Cogn Neurosci 12:711–720.

Grossman ED, Blake R (2002): Brain areas active during visual perception of biological motion. Neuron 35:1167–1175.

Hadjikhani N, Liu AK, Dale A, Cavanagh P, Tootell RBH (1998): Retinotopy and color sensitivity in human visual cortical area V8. Nat Neurosci 1:235–241.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999): "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213–223.

Hamilton R, Pascual-Leone A (1998): Cortical plasticity associated with Braille learning. Trends Cogn Sci 2:168–174.

Hamilton R, Keenan JP, Catala M, Pascual-Leone A (2000): Alexia for Braille following bilateral occipital stroke in an early blind woman. NeuroReport 11:237–240.

Heun R, Freymann K, Erb M, Leube DT, Jessen F, Kircher TT, Grodd W (2006): Successful verbal retrieval in elderly subjects is related to concurrent hippocampal and posterior cingulate activation. Dement Geriatr Cogn Disord 22:165–172.

Hochstein S, Ahissar M (2002): View from the top: Hierarchies and reverse hierarchies in the visual system. Neuron 36:791–804.

Hopfield JJ, Tank DW (1985): "Neural" computation of decisions in optimization problems. Biol Cybern 52:141–152.

Hopfield JJ, Brody CD (2000): What is a moment? "Cortical" sensory integration over a brief interval. Proc Natl Acad Sci USA 97:13919–13924.

Johansson G (1973): Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211.

Johnson SC, Baxter LC, Wilder LS, Pipe JG, Heiserman JE, Prigatano GP (2002): Neural correlates of self-reflection. Brain 125:1808–1814.

Kable JW, Kan IP, Wilson A, Thompson-Schill SL, Chatterjee A (2005): Conceptual representations of action in the lateral temporal cortex. J Cogn Neurosci 17:1855–1870.

Kay KN, Naselaris T, Prenger RJ, Gallant JL (2008): Identifying natural images from human brain activity. Nature 452:352–355.

Kellenbach ML, Brett M, Patterson K (2003): Actions speak louder than functions: The importance of manipulability and action in tool representation. J Cogn Neurosci 15:30–46.

Kiefer M, Sim EJ, Liebich S, Hauk O, Tanaka J (2007): Experience-dependent plasticity of conceptual representations in human sensory-motor areas. J Cogn Neurosci 19:525–542.

Kiefer M, Sim EJ, Herrnberger B, Grothe J, Hoenig K (2008): The sound of concepts: Four markers for a link between auditory and conceptual brain systems. J Neurosci 28:12224–12230.

Kikyo H, Ohki K, Miyashita Y (2002): Neural correlates for feeling-of-knowing: An fMRI parametric analysis. Neuron 36:177–186.

Kim H, Cabeza R (2009): Common and specific brain regions in high- versus low-confidence recognition memory. Brain Res 1282:103–113.

Kim JH, Park KY, Seo SW, Na DL, Chung CS, Lee KH, Kim GM (2007): Reversible verbal and visual memory deficits after left retrosplenial infarction. J Clin Neurol 3:62–66.

Kreifelts B, Ethofer T, Grodd W, Erb M, Wildgruber D (2007): Audiovisual integration of emotional signals in voice and face: An event-related fMRI study. Neuroimage 37:1445–1456.

Lewis JW (2010): Audio-visual perception of everyday natural objects—Hemodynamic studies in humans. In: Marcus J, Naumer PJK, editors. Multisensory Object Perception in the Primate Brain. New York: Springer Science and Business Media, LLC. pp 155–190.

Lewis JW, Phinney RE, Brefczynski-Lewis JA, DeYoe EA (2006): Lefties get it "right" when hearing tool sounds. J Cogn Neurosci 18:1314–1330.

Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA (2005): Distinct cortical pathways for processing tool versus animal sounds. J Neurosci 25:5148–5158.

Lewis JW, Talkington WJ, Puce A, Engel LR, Frum C: Cortical networks representing object categories and perceptual dimensions of familiar real-world action sounds. J Cogn Neurosci (in press).

Lewis JW, Wightman FL, Brefczynski JA, Phinney RE, Binder JR, DeYoe EA (2004): Human brain regions involved in recognizing environmental sounds. Cereb Cortex 14:1008–1021.

Lewis JW, Talkington WJ, Walker NA, Spirou GA, Jajosky A, Frum C, Brefczynski-Lewis JA (2009): Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. J Neurosci 29:2283–2296.

Lissauer H (1890/1988): A case of visual agnosia with a contribution to theory. Cogn Neuropsychol 5:157–192.

Maddock RJ (1999): The retrosplenial cortex and emotion: New insights from functional neuroimaging of the human brain. Trends Neurosci 22:310–316.

Maddock RJ, Garrett AS, Buonocore MH (2001): Remembering familiar people: The posterior cingulate cortex and autobiographical memory retrieval. Neuroscience 104:667–676.

Mahon BZ, Caramazza A (2005): The orchestration of the sensory-motor systems: Clues from neuropsychology. Cogn Neuropsychol 22:480–494.

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.

Martin A (2007): The representation of object concepts in the brain. Annu Rev Psychol 58:25–45.

Mendelsohn A, Furman O, Dudai Y (2010): Signatures of memory: Brain coactivations during retrieval distinguish correct from incorrect recollection. Front Behav Neurosci 4:18.

Montaldi D, Spencer TJ, Roberts N, Mayes AR (2006): The neural system that mediates familiarity memory. Hippocampus 16:504–520.

Moritz S, Glascher J, Sommer T, Buchel C, Braus DF (2006): Neural correlates of memory confidence. Neuroimage 33:1188–1193.

Mutschler I, Schulze-Bonhage A, Glauche V, Demandt E, Speck O, Ball T (2007): A rapid sound-action association effect in human insular cortex. PLoS ONE 2:e259.

Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9:97–113.

Olson IR, Gatenby JC, Gore JC (2002): A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. Brain Res Cogn Brain Res 14:129–138.

Pascual-Leone A, Hamilton R (2001): The metamodal organization of the brain. Prog Brain Res 134:427–445.

Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu WH, 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.

Puce A, Perrett D (2003): Electrophysiology and brain imaging of biological motion. Philos Trans R Soc Lond B Biol Sci 358:435–445.

Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998): Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci 18:2188–2199.

Pylyshyn ZW (2002): Mental imagery: In search of a theory. Behav Brain Sci 25:157–182; discussion 182–237.

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K (2001): Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage 13:669–683.

Ricciardi E, Bonino D, Sani L, Vecchi T, Guazzelli M, Haxby JV, Fadiga L, Pietrini P (2009): Do we really need vision? How blind people "see" the actions of others. J Neurosci 29:9719–9724.

Ries ML, Schmitz TW, Kawahara TN, Torgerson BM, Trivedi MA, Johnson SC (2006): Task-dependent posterior cingulate activation in mild cognitive impairment. Neuroimage 29:485–492.

Ries ML, Jabbar BM, Schmitz TW, Trivedi MA, Gleason CE, Carlsson CM, Rowley HA, Asthana S, Johnson SC (2007): Anosognosia in mild cognitive impairment: Relationship to activation of cortical midline structures involved in self-appraisal. J Int Neuropsychol Soc 13:450–461.

Rizzolatti G, Craighero L (2004): The mirror-neuron system. Annu Rev Neurosci 27:169–192.

Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996): Premotor cortex and the recognition of motor actions. Brain Res Cogn Brain Res 3:131–141.

Robins DL, Hunyadi E, Schultz RT (2009): Superior temporal activation in response to dynamic audio-visual emotional cues. Brain Cogn 69:269–278.

Rockland KS, Ojima H (2003): Multisensory convergence in calcarine visual areas in macaque monkey. Int J Psychophysiol 50:19–26.

Roder B, Rosler F (2003): Memory for environmental sounds in sighted, congenitally blind and late blind adults: Evidence for cross-modal compensation. Int J Psychophysiol 50:27–39.

Roder B, Stock O, Bien S, Neville H, Rosler F (2002): Speech processing activates visual cortex in congenitally blind humans. Eur J Neurosci 16:930–936.

Sadato N, Okada T, Honda M, Yonekura Y (2002): Critical period for cross-modal plasticity in blind humans: A functional MRI study. Neuroimage 16:389–400.

Sadato N, Pascual-Leone A, Grafman J, Ibanez V, Deiber MP, Dold G, Hallett M (1996): Activation of the primary visual cortex by Braille reading in blind subjects. Nature 380:526–528.

Safford AS, Hussey EA, Parasuraman R, Thompson JC (2010): Object-based attentional modulation of biological motion processing: Spatiotemporal dynamics using functional magnetic resonance imaging and electroencephalography. J Neurosci 30:9064–9073.

Schmitz TW, Rowley HA, Kawahara TN, Johnson SC (2006): Neural correlates of self-evaluative accuracy after traumatic brain injury. Neuropsychologia 44:762–773.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995): Borders of multiple visual areas in humans revealed by functional MRI. Science 268:889–893.

Shipp S, Watson JD, Frackowiak RS, Zeki S (1995): Retinotopic maps in human prestriate visual cortex: The demarcation of areas V2 and V3. Neuroimage 2:125–132.

Simmons WK, Reddish M, Bellgowan PS, Martin A (2009): The selectivity and functional connectivity of the anterior temporal lobes. Cereb Cortex 20:813–825.

Simons JS, Peers PV, Mazuz YS, Berryhill ME, Olson IR (2010): Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. Cereb Cortex 20:479–485.

Slotnick SD, Thompson WL, Kosslyn SM (2005): Visual mental imagery induces retinotopically organized activation of early visual areas. Cereb Cortex 15:1570–1583.

Stevens AA, Snodgrass M, Schwartz D, Weaver K (2007): Preparatory activity in occipital cortex in early blind humans predicts auditory perceptual performance. J Neurosci 27:10734–10741.

Stevenson RA, James TW (2009): Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. Neuroimage 44:1210–1223.

Talairach J, Tournoux P (1988): Co-Planar Stereotaxic Atlas of the Human Brain. New York: Thieme Medical Publishers.

Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM (2004): Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. J Neurophysiol 91:1282–1296.

Taylor KI, Stamatakis EA, Tyler LK (2009): Crossmodal integration of object features: Voxel-based correlations in brain-damaged patients. Brain 132:671–683.

Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006): Binding crossmodal object features in perirhinal cortex. Proc Natl Acad Sci USA 103:8239–8244.

Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, Fazio F, Rizzolatti G, Cappa SF, Perani D (2005): Listening to action-related sentences activates fronto-parietal motor circuits. J Cogn Neurosci 17:273–281.

Tootell RBH, Hadjikhani NK, Mendola JD, Marrett S, Dale AM (1998): From retinotopy to recognition: fMRI in human visual cortex. Trends Cogn Sci 2:174–183.

Valenstein E, Bowers D, Verfaellie M, Heilman KM, Day A, Watson RT (1987): Retrosplenial amnesia. Brain 110 (Part 6):1631–1646.

Van Essen DC (2005): A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. Neuroimage 28:635–662.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001): An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc 8:443–459.

Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005): Parietal lobe contributions to episodic memory retrieval. Trends Cogn Sci 9:445–453.

Wakusawa K, Sugiura M, Sassa Y, Jeong H, Horie K, Sato S, Yokoyama H, Tsuchiya S, Kawashima R (2009): Neural correlates of processing situational relationships between a part and the whole: An fMRI study. Neuroimage 48:486–496.

Wakusawa K, Sugiura M, Sassa Y, Jeong H, Horie K, Sato S, Yokoyama H, Tsuchiya S, Inuma K, Kawashima R (2007): Comprehension of implicit meanings in social situations involving irony: A functional MRI study. Neuroimage 37:1417–1426.

Walter H, Adenzato M, Ciaramidaro A, Enrici I, Pia L, Bara BG (2004): Understanding intentions in social interaction: The role of the anterior paracingulate cortex. J Cogn Neurosci 16:1854–1863.

Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009): Natural scene categories revealed in distributed patterns of activity in the human brain. J Neurosci 29:10573–10581.

Warrington EK, Shallice T (1984): Category specific semantic impairments. Brain 107(Part 3):829–854.

Weaver KE, Stevens AA (2007): Attention and sensory interactions within the occipital cortex in the early blind: An fMRI study. J Cogn Neurosci 19:315–330.

Yonelinas AP, Otten LJ, Shaw KN, Rugg MD (2005): Separating the brain regions involved in recollection and familiarity in recognition memory. J Neurosci 25:3002–3008.

# APPENDIX

List of the sound stimuli judged to be directly produced by human actions (61 of 105 stimulus pairs presented). Numbers in parentheses indicate percentage of EB and SC individuals that both recognized the forward-played sounds (RF) and did not recognized backward-played sounds (NB), and thus retained for the main analysis. The inset charts the number of forward- plus backward-played sound pairs retained for analysis for the EB group (black circles) and the SC group (gray squares). Refer to text for other details.

**Scan 1**
ice dropped into glass (50,64)
glass breaking #1 (60,50)
rotary phone dialing (70,57)
hammer hitting anvil (40,71)
audience clapping (80,71)
footsteps (50,43)
doorbell (50,29)
hammer hitting nail #1 (80,71)
glass breaking #2 (60,86)

**Scan 2**
ping pong rally (60,79)
card shuffle (20,57)
scissors cutting hair (20,21)
keys jingle & toss (10,36)
machine gun (70,79)
drum accent (60,79)
golf putt into cup #1 (80,50)

**Scan 3**
tennis rally (60,71)
typing on computer (50,57)
cracking an egg (30,50)
bowling strike #1 (70,79)
toasting with glass (70,71)
beating eggs (50,50)
car crash (20,14)
slot machine payout (10,64)
door open & close (90,93)
pouring beer (30,21)

**Scan 4**
billiards break shot (80,86)
revolver gun shot (70,93)
draining bathtub (50,86)
toaster (40,64)
removing a cork (50,64)
manual typewriter #1 (40,29)
whisky pouring into glass (70,86)
hammering nail #2 (60,86)
opening can of soda (0,21)
knocking on door (80,93)

**Scan 5**
small gun fire (60,71)
ping pong rally (70,79)
pool hall shot (80,79)
manual typewriter #2 (90,50)
explosion (70,50)
ascending chimes (60,79)
bongo drums (80,86)
chopping carrot (10,14)
stapling paper (60,57)

**Scan 6**
push in coin slot (30,64)
manual typewriter #3 (40,21)
dribbling basketball (50,79)
water dripping in metal dish (60,71)
toilet flushing (0,43)
paper cutter cutting (20,21)
bowling strike #2 (80,86)
billiards shot into pocket (80,79)
dive into swimming pool (20,64)
golf putt into cup #2 (70,71)l

**Scan 7**
filling bathtub (50,64)
bite & chew chips (10,36)
basketball thru net & bounce (70,57)
metal file cabinet closing (50,71)
racquetball rally (80,86)
burp (40,71)