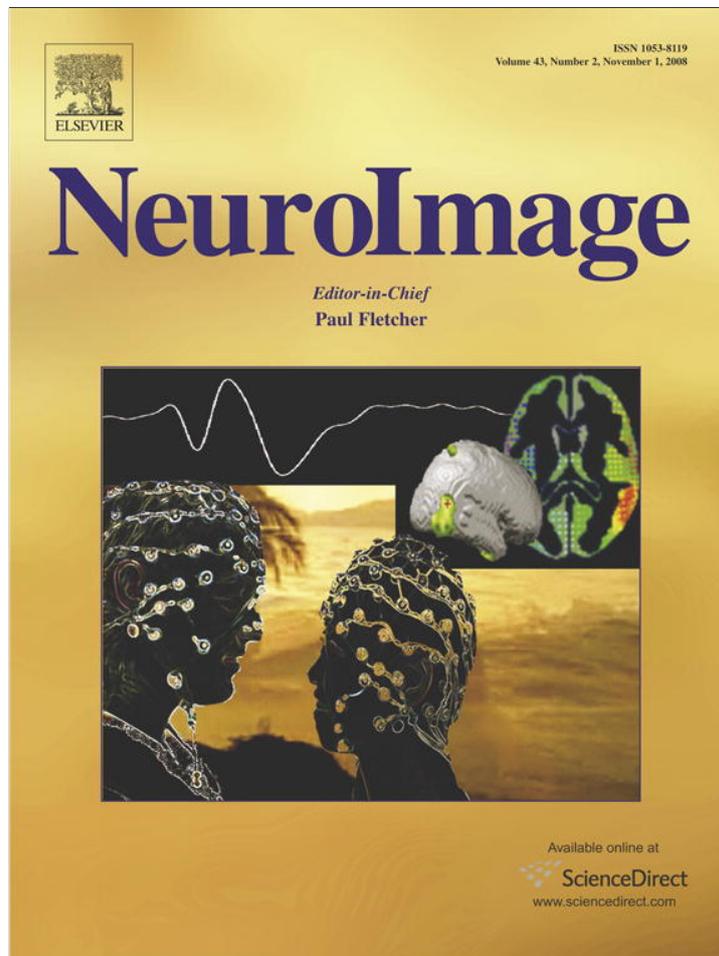


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

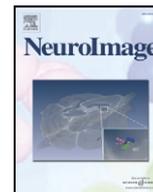
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Crinkling and crumpling: An auditory fMRI study of material properties

Stephen R. Arnott^{a,*}, Jonathan S. Cant^{a,b}, Gordon N. Dutton^{c,d}, Melvyn A. Goodale^{a,b}

^a CIHR Group for Action and Perception, Department of Psychology, University of Western Ontario, London, Ontario, N6A 5C2 Canada

^b Neuroscience Program, University of Western Ontario, London, N6A 5C2 Canada

^c Royal Hospital for Sick Children, Yorkhill, Glasgow, Scotland G3 8SJ, UK

^d Glasgow Caledonian University, Cowcaddens Road, Glasgow, Scotland G4 0AB, UK

ARTICLE INFO

Article history:

Received 21 December 2007

Revised 25 May 2008

Accepted 12 July 2008

Available online 29 July 2008

Keywords:

Blind

Parahippocampus

Parietal

Sound

Texture

Visual

ABSTRACT

Knowledge of an object's material composition (i.e., what it is made of) alters how we interact with that object. Seeing the bright glint or hearing the metallic crinkle of a foil plate for example, confers information about that object before we have even touched it. Recent research indicates that the medial aspect of the ventral visual pathway is sensitive to the surface properties of objects. In the present functional magnetic resonance imaging (fMRI) study, we investigated whether the ventral pathway is also sensitive to material properties derived from sound alone. Relative to scrambled material sounds and non-verbal human vocalizations, audio recordings of materials being manipulated (i.e., crumpled) in someone's hands elicited greater BOLD activity in the right parahippocampal cortex of neurologically intact listeners, as well as a cortically blind participant. Additional left inferior parietal lobe activity was also observed in the neurologically intact group. Taken together, these results support a ventro-medial pathway that is specialized for processing the material properties of objects, and suggest that there are sub-regions within this pathway that subserve the processing of acoustically-derived information about material composition.

© 2008 Elsevier Inc. All rights reserved.

Introduction

In many cases, the mere glimpse of an object's texture, colour or shape is all that is needed to identify it. The neural correlates underlying this ability have been readily accounted for by a vision-for-perception pathway in which information from striate cortex projects to ventral brain areas in the inferotemporal cortex (Goodale and Milner, 1992). Whereas lateral areas of this pathway have been shown to be especially sensitive for identifying objects based on shape (James et al., 2003; Malach et al., 1995), research suggests that the medial aspect of this pathway is better suited for perceiving an object's material properties (e.g., texture, material composition. Cant and Goodale, 2007; Peuskens et al., 2004; Puce et al., 1996). Thus, when functional magnetic resonance imaging (fMRI) is conducted on participants viewing pictures of 3D non-sense objects that vary in material composition, shape, or colour, attention to material differences elicits relatively greater haemodynamic activity along the right collateral sulcus and inferior occipital gyrus as well as the left lingual sulcus and inferior temporal sulcus (Cant and Goodale, 2007). This, in context with other findings showing that damage to the lateral occipital complex (LOC) affects the ability to identify objects from line drawings but not from photographs and drawings with texture or colour details added (James et al., 2003; Milner et al., 1991), supports

the claim that the medial part of the ventral pathway is specialized for processing material-property information.

Yet information about an object's material properties can also be acquired through non-visual routes. For example, just as the sight of the bright highlights on a foil plate confers some information about that object's aluminium composition, so too does the sound of its metallic crinkling convey a similar type of information, even when that foil plate is hidden from view. The purpose of the current study was to investigate whether material-processing information derived from sound (i.e., non-visual routes) would also activate the same or very proximal regions of the ventral stream that have been shown to be sensitive to visual texture.

Although the study of acoustically defined material-property information has received less research than its visual counterpart (but see Giordano and McAdams, 2006; Klatzky et al., 2000; Wildes and Richards, 1988), there are reasons to expect functional crossover. For one, as is the case with vision, there are auditory data from both animal (Rauschecker and Tian, 2000; Kaas and Hackett, 1999; Romanski et al., 1999) and human (Alain et al., 2001; Arnott et al., 2004; Arnott et al., 2005; Clarke et al., 2002; Rämä et al., 2004) studies supporting a 'ventral' pathway that is specialized for perceiving and identifying objects/events in the environment. This pathway comprises areas along the superior temporal gyrus (STG); recruiting more rostral regions with increasing sound complexity), extending into ventral prefrontal cortex (Alain et al., 2001; Rämä et al., 2004; Arnott et al., 2004). In addition to these more traditional auditory areas, there

* Corresponding author. Fax: +1 519 661 3961.

E-mail address: sarnott3@uwo.ca (S.R. Arnott).

is increasing evidence that this pathway may interact with ventral visual brain areas. For example, several fMRI studies (Arnott et al., 2008; von Kriegstein et al., 2006; von Kriegstein et al., 2005) have demonstrated that when listeners attend to the identity of voices, haemodynamic increases are found not only in the temporal pole but in regions proximal to the well-documented visual fusiform face area (FFA, Kanwisher et al., 1997; Grill-Spector et al., 2004; Kanwisher et al., 1997; Grill-Spector et al., 2004). Indeed, some research suggests that auditory information can exert influences even earlier in the visual hierarchy (Arnott and Goodale, 2006; Budinger et al., 2006; Falchier et al., 2002).

Still another reason to expect areas in the ventral medial pathway to respond to auditory material-property information stems from the fact that knowledge of auditory material-property information is associative. That is, because there is nothing innately identifiable about acoustic material properties, the ability to identify the sound of say, crinkling aluminium foil, is predominately derived from experiences in which an observer hears the sound that is made as he or she observes a foil object being manipulated. In other words, knowledge about the acoustic properties of various materials comes about from an accrual of many multi-modal experiences. Accordingly, one might expect areas in and/or around the region showing sensitivity to visual surface properties to also process similar auditory-derived information. The anterior portion of the ventral medial visual pathway in particular, is closely situated to regions of the medial temporal lobe that are known to support mnemonic and memory processes (e.g., Buckley, 2005). Interestingly, fMRI data have demonstrated a significant amount of spatial overlap between cortical areas involved in processing material information derived from the visual domain, and those areas involved in the encoding and recognition of scenes (i.e., the parahippocampal place area (PPA; Epstein and Kanwisher, 1998), especially in the right hemisphere (Cant and Goodale, 2007). Moreover, it has been proposed that processing of scenes and contextual information in the PPA owes as much to memory as it does to bottom-up visual input (Fenske et al., 2006; Bar, 2004). Finally, there are direct pathways connecting auditory cortex with areas in the medial temporal lobe. Retrograde staining of monkey parahippocampus, for example, has revealed projections from auditory areas in the superior temporal gyrus (Lavenex et al., 2002, 2004; Suzuki and Amaral, 1994).

In the present experiment, we first localized the area in the medial ventral stream most strongly associated with the visual processing of surface properties of objects. We then constructed a region of interest (ROI) that encompassed this area and the surrounding perimeter. We reasoned that any auditory areas associated with the material properties of objects would be located around the same region by virtue of the fact that there should be multiple sensory routes to the same material-property representations.

In addition to investigating acoustic material processing in neurologically intact participants, we tested this ability in a cortically blind patient who, with the exception of a spared right parahippocampus and a small tag of tissue in the anterior calcarine sulcus, had suffered bilateral lesions to much of her occipital cortex. We expected that even though she was unlikely to show any brain activation associated with the visual processing of surface properties, she might still show activation to sounds associated with the material properties of objects.

Materials and methods

Participants

Though fourteen volunteers participated in this experiment, neuroimaging data from only the top nine performers were analyzed (see Results; mean age=26 years; standard deviation=5.2 years; all right-handed; normal hearing and normal or corrected-to-normal vision; 5 males). Additionally, one patient (reported below) also

participated. Neurologically intact participants were members of the University of Western Ontario community and included undergraduates, graduates and postdoctoral fellows experienced at keeping still and maintaining fixation during fMRI experiments. Before participating, the volunteers gave their informed consent in accordance with the Declaration of Helsinki. The procedures and protocols were approved by the Review Board for Health Sciences Research Involving Human Participants for the University of Western Ontario and the Robarts Research Institute (London, Ontario, Canada).

Auditory stimuli

The audio stimuli were recorded in a soundproof chamber using an omni-directional microphone (Sony ECM-T115) situated approximately 40 cm in front of the sound source. Two categories of natural sounds were recorded: various materials being crumpled in someone's hands, and human non-verbal vocalizations. For the latter sounds, four male volunteers were assigned to one of four non-verbal vocalizations (coughing, snoring, throat clearing, or yawning) and asked to produce at least twelve instances of that sound. Material sounds were created by having a person crumple (i.e., crush) various materials between his left and right hands, at a moderate and consistent speed (approximately 4–5 seconds (s) in length). The four types of materials used were sheets of paper (standard 22×28 cm photocopy sheets, 75 g/m²), foil (31×22 cm heavy duty strength aluminium foil, 0.015 cm thick), styrofoam (15×13 cm extruded polystyrene, 0.6 cm thickness), and plastic (23×15 cm snack food bag composed of two layers of oriented polypropylene bonded with a layer of low density polyethylene). With the exception of the plastic sounds in which the same polypropylene bag was repeatedly used, each sound was created from a 'new' intact sheet. Twenty manipulations of each material were recorded. Post-recording editing (Adobe Audition, V1.0) consisted of trimming each audio clip to two seconds and down-sampling each 66 kHz, 32-bit resolution, mono sound file to a rate of 48 kHz with 16-bit resolution. A third category of sounds (i.e., scrambled material) was created by randomizing the amplitude envelopes of each material sound stimulus. Specifically, the intensity envelopes for each two-second wave file were segmented into nineteen 105 ms time windows and then randomly reordered (Matlab version 6.5.0, The MathWorks, Inc.) and reapplied to the waveform. Consequently, the resulting waveforms no longer possessed the overall temporal envelope of a crinkling material sound, but sounded more like band-passed noise.

Following the acoustic processing, eight of the 'best' exemplars of each sound were selected from the human and material categories for use in the actual experiment. While the human sounds were selected based on experimenter preference, material sound selection was based on a pilot experiment where 80 sounds (20 each of paper, foil, styrofoam, and plastic) were randomly played (Superlab, version 2.0.4; Cedrus Corporation, San Pedro, CA) to seven listeners (none of whom participated in the subsequent fMRI experiment) over fMRI-compatible circumaural headphones. Participants were asked to identify these material sounds with a button press as quickly and as accurately as possible. Based on mean accuracy, the top eight sound clips within each of the material categories were selected for use in the fMRI experiment. The scrambled versions of each of these material sounds were selected as the scrambled material stimuli. As a final step, all sound files to be used in the experiment (i.e., material, scrambled material, and human sounds) underwent a group-waveform normalization procedure and were then amplitude-gated with a 10 ms onset/offset sine ramp to prevent clicks.

Visual stimuli

The visual experiment was designed to address a separate question (Cant et al., submitted for publication), but was incorporated in the

same experimental session as the auditory task so that one of its manipulations could serve as the basis for a creating a region of interest. The stimuli have been extensively described elsewhere (Cant and Goodale, 2007). Each stimulus was a 640×480 pixels digitized (Discreet 3DS Max, Montreal, Quebec, Canada) unfamiliar, bilaterally symmetrical non-sense object adapted from stimuli used in an earlier series of studies (Harman et al., 1999; Humphrey and Khan, 1992). Four object forms were used, each of which was rendered in one of four materials (oak, metal, marble, and aluminium foil) and four hues (red, yellow, blue, and green), giving a total of sixty-four objects. For the purposes of the present study, only the material category manipulation was relevant.

Experimental design

Auditory experiment

Prior to the fMRI session, each participant completed a brief behavioural task in order to determine *a priori* whether he or she could identify the material sounds (at least in a non-MRI setting). For this experiment, 20 scrambled material and 80 material (i.e., 20 each of styrofoam, paper, foil and plastic) sounds were randomly presented to the listener over headphones. Participants were told that they would hear plastic, styrofoam, paper, aluminium foil, or scrambled versions of these and that after the onset of each sound they were to press a button indicating which of the five sounds they had heard. None of the listeners was informed as to how these material sounds were obtained. Those who achieved a mean accuracy of at least 75% for each material category were invited to participate in an fMRI session.

For the fMRI session, digitized sounds were converted to an air signal by a pneumatic, non-magnetic transducer in the MRI room and sent through two channels (3 meter long, 19 oz flexible plastic tubing; approximately 1.2 cm in diameter) to acoustically padded (30 dB attenuation) circumaural, fMRI-compatible headphones (Resonance Technology, Northridge, CA) at an intensity that each participant judged to be comfortably audible over the background scanner noise. As with the visual localizer, participants were asked to fixate on a central cross throughout the experiment. Each trial of this slow-event related design began with a 300 ms, 1000 Hz alerting signal that was followed 700 ms later by one of the three types of random sound clips (Material, Scrambled Material, or Human). Participants were asked to listen to the entire duration of each clip and then indicate which category the sound belonged to by pressing one of three buttons.

Note, unlike the behavioural session, the fMRI experiment did not require that the listener explicitly identify the specific within-category sound but merely categorize it as a Material, Scrambled Material or Human sound. Nevertheless, most participants reported at the end of the experimental session that they had been able to identify and discriminate the different types of sounds. Those who reported difficulty typically confused foil, plastic, and scrambled material. This difficulty may relate to masking created by the scanner noise (Healy et al., 2007), as well as high frequency filtering effects of the plastic tubing delivering the sound into the headphones. Following sound offset, there was a 14.5 s silent period to allow the blood oxygenation level dependent (BOLD) response to return to baseline levels before the next trial began. Each block contained 18 trials and took approximately 5:15 min to complete. All participants completed a minimum of five blocks with most completing six.

Visual experiment

As mentioned earlier, although the experiment was designed to address a separate hypothesis (Cant et al., submitted for publication), it was run in the same imaging session so that its 'material-properties' manipulation could serve as a functional localizer for the auditory material imaging data. This visual experiment was run prior to the auditory experiment. Each image was projected via an LCD projector

(NEC VT540, Japan, screen resolution of 800×600) onto a screen mounted above the participant's waist as he or she lay in the bore of the magnet. The participant viewed the image through a mirror angled 45° from the brow-line, directly above the eyes. Distance from the participant's eyes, via the mirror-to-the-screen was ~60 cm. Prior to entering the magnet, participants viewed images of the four materials and verbal descriptions of each material were provided by the experimenters (i.e. "this is metal, this is wood", etc.).

In each run, participants were presented with 16 s blocks of trials where either the material changed from trial-to-trial (form and colour were held constant), only the colour changed, or only the form changed. Only the material changing blocks were of interest to the present study. Each of these blocks was interleaved with a 16 s block where no feature changed (i.e., every stimulus was presented in the same material, colour and form). The participant's task throughout the entire experiment was to attend to the images while keeping their eyes fixated on the fixation cross.

A single trial in each experimental block consisted of an 800 ms image presentation, followed by a 200 ms blank screen presentation. Thus, each of the 16 trials in each block lasted for a duration of 1 s. Three material change blocks were presented within each run. All participants completed a minimum of four experimental runs, with some completing five, depending on time constraints.

Data acquisition

Scans were conducted using a 4-T Siemens-Varian whole-body MR scanner (Erlangen Germany; Palo Alto, CA) with a standard cylindrical quadrature head coil. Participants lay in a supine position with a piece of surgical tape placed across their forehead and affixed to either side of the MRI bed to help minimize head motion. Functional imaging was performed to measure the BOLD effect with optimal signal contrast. For the functional scans in the visual experiment, 17 contiguous axial slices were acquired using a navigator echo-corrected, slice-interleaved multi-shot (2-shot), T2*-weighted, echo-planar imaging pulse sequence (volume acquisition time=2000 ms, TR=1000 ms, TE=15.0 ms, flip angle 40°, field of view 22.0 cm, 3.4 mm×3.4 mm×5.0 mm voxel size, 64×64 effective acquisition matrix). With the exception of 21 slices being obtained using a volume acquisition time of 2500 ms, TR=1250 ms and a flip angle of 45°, the functional scans in the auditory experiment were the same as that in the visual. All functional scans were aligned to a high-resolution volumetric anatomical MRI that was acquired following the visual and auditory experiments. The anatomical run consisted of single-shot T1-weighted images collected with the same slice orientation (3D magnetization-prepared turbo FLASH acquisition with inversion time (T1)=750 ms, TR=45 ms, TE=3.0 ms, 22.0 cm×22.0 cm field of view, 256×256 acquisition matrix, 112 axial slices, 0.86 mm×0.86 mm×1.25 mm voxel size).

fMRI analysis

Data processing and analyses were performed using Brain Voyager QX software (version 1.7, Brain Innovation, Maastricht, The Netherlands). After undergoing linear trend removal, high-pass filtering (3 cycles) and a correction for serial correlations (using a Brain Voyager script that adjusts the standard error of the beta weights so as to avoid the problem of inflated *t*-values associated with the autocorrelation of neighbouring measurement time points), functional scans were co-registered to the last functional run before the anatomical volume and then transformed into a common stereotaxic space (Talairach and Tournoux, 1988). The functional data were neither spatially smoothed nor, due to the multi-shot acquisition technique, subjected to a slice scan time correction. Prior to any pre-processing, time-course movies of the individual runs for each participant were examined and quantitatively evaluated using Brain

Voyager's 3D motion correction algorithm to ensure that there was no head motion that exceeded 1 mm in translation and/or 1° in rotation. Functional runs that exceeded these parameters were excluded from further analysis. These precautions, in addition to the piece of surgical tape affixed to each participant's forehead, ensured that the data were not contaminated by motion.

For the visual localizer task, baseline haemodynamic activity was taken as the 16 s of fixation between each stimulus block, and separate predictors for each stimulus block were created (Material Change, Form Change, Colour Change, and No Change). The reference haemodynamic-related function (HRF) for both the visual and auditory task was defined by Brain Voyager's two gamma haemodynamic response function (onset of curve=0 ms, time to response peak=5 s, response dispersion=1, undershoot ratio=6, time to undershoot peak=15 s, and undershoot dispersion of 1). For the auditory task, general linear model (GLM) analyses were performed with separate predictors for each stimulus (i.e., correctly identified material sounds (*M*), correctly identified scrambled material (*X*), and correctly identified human sounds (*H*)) as well as for the cue tone at the beginning of each trial. Incorrectly identified sounds were grouped as a condition of no interest and were omitted from further analysis. Finally, to ensure that the GLM findings would not be skewed by strong effects exerted by a subset of the participants or by an unequal number of runs from participants, data from each single participant were grouped (i.e., the predictors of all runs belonging to that participant were pooled) using the 'separate subject predictors' function in Brain Voyager.

ROI analysis

Statistical analysis of the fMRI data was a multi-step process. First, the visual experimental data were used to create broad 'material-processing' regions of interest (ROIs). This was accomplished by means of a simple contrast of the Material Change (vM) condition with the No Change (vN) condition (i.e., vM – vN; $p < 0.00001$). Under the premise that auditory material processing would occur in the same or very proximal brain areas recruited for visual material processing, we then expanded these clusters in every direction until they formed two large ROIs, one in the left and one in the right ventro-medial cortex. These 'material-processing' ROIs were then used to investigate the fMRI data from the auditory experiment. For each ROI, a spatial cluster extent threshold was calculated (1000 Monte Carlo simulations) using AlphaSim (Cox, 1996) to assess how large a cluster of activation had to be in a given ROI in order to be considered significant at a particular threshold value. The second step of analysis, involved analyzing the fMRI data from the auditory experiment by means of a conjunction contrast of Material with Scrambled Material and Human sounds (i.e., $M - X \cap M - H$). Any auditory-elicited 'material' activations occurring both inside and outside of the visual material ROIs were noted and quantified. Note that the type of conjunction contrast implemented in these Brain Voyager analyses was of the type recommended by Nichols et al. (2005) in which the minimum statistic is compared to the conjunction null (as opposed to the global null).

As an additional means of comparing functional activity, event-related haemodynamic response plots were created in some instances. In such case, percent signal change values were extracted from activation clusters of interest using BV's file-based procedure and a baseline of 2. TRs prior to stimulus onset.

Voxel-wise analysis

As a final analysis, a voxel-wise GLM was carried out. For this analysis, the total number of functional brain voxels was calculated by overlaying Talairach functional volumes from all participants, setting a minimum threshold of overlap (i.e., BV segmentation value of 140), quantifying this voxel total, and converting it to voxel units of

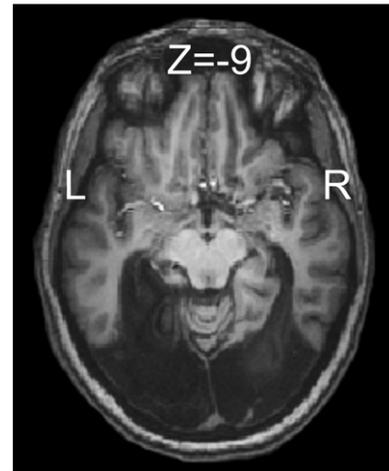


Fig. 1. Anatomical T1-weighted axial MRI image of patient MC (Z = -9 mm).

1 mm × 1 mm × 1 mm. This procedure calculated the whole-brain total to be of 985,720 mm³. Based on this number, it was determined, using AlphaSim, that at uncorrected GLM thresholds of $p < 0.001$, the minimum cluster size needed to achieve 95% reliability was 81 voxels.

Cortically blind participant

To further examine the role that vision plays in the perception of 'material' sounds, a cortically blind 38 year-old right-handed female (MC) also participated in the experiments. MC's injuries were incurred eight years earlier as a result of severe respiratory infection and subsequent hypotension (Dutton, 2003) that precipitated a hypoxic ischemic episode. Apart from some sparing in the anterior right calcarine cortex and right parahippocampus, the resulting cortical damage encompassed her occipital cortex bilaterally; encroaching dorsally into posterior portions of the temporal lobes bilaterally, as well as the right posterior parietal lobe (see Fig. 1). As reported previously (Dutton, 2003), MC possesses some preserved visual motion processing. That said, MC is a high-functioning, motivated individual with good insight into her deficits. Audiological testing revealed normal pure tone thresholds and normal middle ear function as assessed by tympanometric measurements.

Prior to the fMRI experiment, MC was given a brief behavioural test to determine whether she, like the neurologically intact people, could identify the various material sounds. The fMRI experimental session was carried out two days later. Due to the fact that she had participated in a host of other fMRI and behavioural experiments in the hours and days prior to the fMRI testing, MC was able to complete four experimental runs of the auditory fMRI task before she got tired. Data processing was therefore the same as outlined above.

Results

The auditory task was straightforward and nine of the fourteen participants performed almost flawlessly. Paired *t*-tests performed on these participants' mean accuracy data did not show any significant differences between the three sound categories ($M = 0.98$, $X = 0.96$, $H = 0.98$; all p values > 0.1). The five remaining participants were omitted from the fMRI analyses because they could not differentiate between material sounds and scrambled versions of those sounds in the magnet. Thus, even though these five participants could correctly categorize material or humans sounds almost perfectly ($M = 0.95$, $H = 0.97$, $p > 0.4$), they often mis-categorized scrambled sounds as real material sounds (mean X accuracy = 0.67). Paired *t*-tests carried out on the accuracy data for these five participants confirmed this relative difficulty (M vs. X , $t = 3.99$, $p = 0.01$; H vs. X , $t = 4.24$, $p < 0.01$). Possible

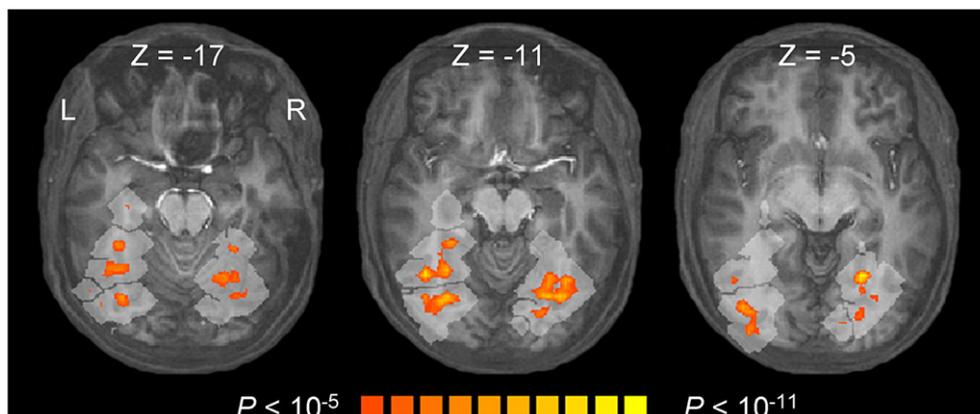


Fig. 2. Visual ‘material-processing’ haemodynamic activity ($N=9$). Material changing blocks (vM) > No change blocks (vN). The two light gray areas denote regions of interest (ROIs) created by enlarging the borders of the haemodynamic activity by 10 mm^3 .

explanations for this performance include inattention as well as stimulus degradation due to high frequency filtering from the plastic tubing and the acoustic masking properties of the MRI scanner (Healy et al., 2007; Menon et al., 2002).

Visual localizer experiment

In keeping with previous investigations of cortical areas involved in the visual processing of surface properties of objects (Cant and Goodale, 2007), a GLM voxel-wise analysis ($N=9$) of vM – vN ($p=0.00001$, uncorrected), revealed two extents of material-specific activation in the right and left ventro-medial areas (Fig. 2). These hemispheric activations encompassed areas of the fusiform gyrus, lingual sulcus, collateral sulcus and posterior regions of parahippocampal cortex (with peak activations located in bilateral medial occipital and fusiform gyri, left inferior temporal gyrus and left parahippocampal gyrus; see Table 1). These ventro-medial clusters were then enlarged in every direction until they formed one large ROI in each hemisphere (right hemisphere=44,346 voxels; left hemisphere=56,538 voxels). The amount of enlargement was 10 cubic voxels. These two ROIs were then used to probe the results of the auditory experiment. AlphaSim calculations determined that auditory activation clusters of 108 and 134 voxels within the right and left ventro-medial areas, respectively, could be considered significant at thresholds of $p=0.01$.

Auditory experiment

To determine whether perceiving material properties through sound alone recruited brain areas located in the visual material ROIs, a

Table 1
Visual localizer results: peak activations associated with Material Change > No Change

Location	BA	Peak t value	X	Y	Z	Number of voxels; $t_{4939} > 5.00$
Left medial occipital gyrus	19	5.91	-42	-76	-5	267
Left inferior temporal gyrus	37	7.28	-42	-58	-14	532
Left fusiform gyrus	18	5.63	-36	-85	-2	125
Left fusiform gyrus	18	7.27	-33	-73	-14	444
Left parahippocampal gyrus	35	6.12	-27	-40	-11	182
Left medial frontal gyrus	46	5.59	-33	26	25	73
Left superior parietal lobule and medial occipital gyrus	7/19	6.39	-27	-70	34	141
Right fusiform gyrus	37/19	7.52	24	-58	-5	1186
Right fusiform gyrus	19	5.86	24	-79	-8	72
Right medial occipital gyrus	19	5.57	27	-79	13	111

BA = Brodmann's area.

‘material sound’ conjunction analysis was carried out. This involved testing for voxels with correctly identified material BOLD activity that was significantly greater than that associated with correctly identified scrambled material sounds as well as correctly identified human sounds (i.e., $M-X \cap M-H$). At a threshold criterion of $p=0.01$, such an auditory activation cluster was found in the right posterior parahippocampus, with the peak activation occurring at the very rostral medial edge of the visual ROI ($X=19, Y=-41, Z=-11$; see Fig. 3). Because of its 109 voxel size, it satisfied the minimum cluster size parameters set by AlphaSim. There was a small area of auditory activity in the left parahippocampal gyrus ($X=-12, Y=-34, Z=-14$) as well, but its size was below the minimum set by AlphaSim. Further exploration of the right parahippocampal auditory area suggested that it occupied a different foci from the areas activated by the visual material localizer experiment in the sense that when the visual material contrast (i.e., vM – vN) map was replotted at a very lax statistical threshold of $p < 0.05$ (uncorrected), the statistical map failed to entirely encompass the auditory material area and did not even encapsulate the peak voxel of the auditory activation (see inset of Fig. 3).

The only significant area of activation (i.e., >81 voxels at $p < 0.001$) resulting from a voxel-wise search using the same conjunction analysis of material sound (i.e., $M-X \cap M-H$) was located in the left inferior parietal lobule ($X=-45, Y=-47, Z=39$; 179 voxels; Fig. 4). This area appeared to be localized in the vicinity of anterior intraparietal area (i.e., area AIP), a region well-documented in the macaque to be involved in visually-guided grasping movements towards objects (for a review, see Culham et al., 2006; Culham and Valyear, 2006). Consistent results have shown that the human homologue to macaque AIP (hAIP) is located at the inferior post-central sulcus (iPS) and anterior intraparietal sulcus (aIPS) junction (Binkofski et al., 1998; Culham et al., 2003, 2006; Culham and Valyear, 2006; Faillenot et al., 1997; Frey et al., 2005). Accordingly this anatomical junction was localized in each participant and the corresponding Talairach coordinates were averaged across the nine participants, giving an estimated left hAIP location of $X=-41$ ($SD=2.7$), $Y=-37$ (2.7), $Z=41$ (4.7). This hAIP location was bolstered when the Talairach coordinates from several hAIP neuroimaging studies (as reported in Fig. 3 of (Castiello and Begliomini, 2008)) were averaged together, providing a mean left hAIP location of $X=-39$ ($SD=4.8$), $Y=-40$ (5.2), $Z=34$ (28.3). As can be seen in Fig. 4, the auditory material-processing area was situated caudal to both hAIP estimates; 10 mm behind the average junction of the iPS and aIPS, and 7 mm behind the hAIP location obtained by averaging the coordinates from the hAIP meta-analysis (Castiello and Begliomini, 2008). Finally, the auditory IPL area was converted to an ROI in order

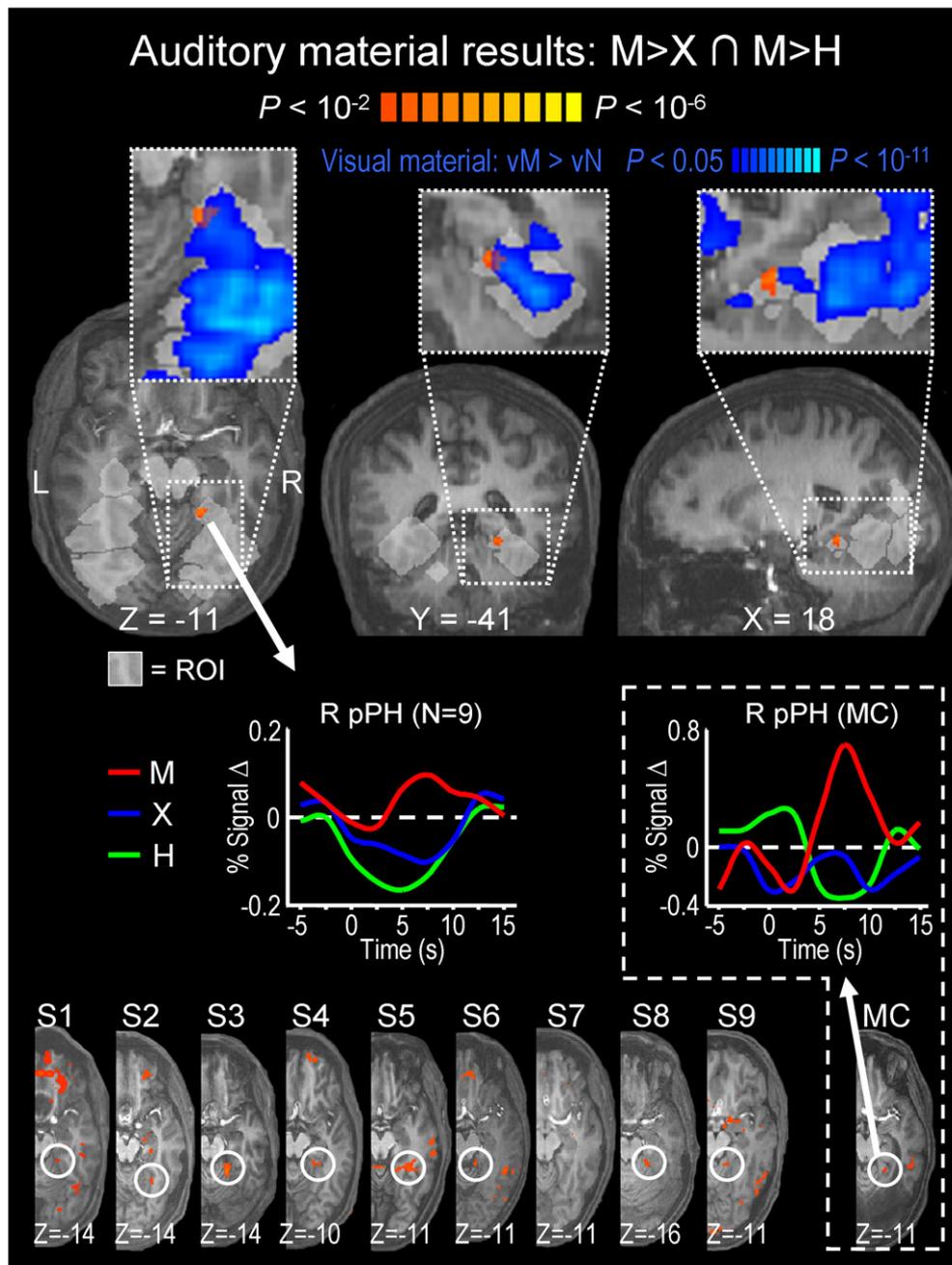


Fig. 3. Auditory ‘material-processing’ conjunction results from the ROI analysis ($N=9$). Magnified areas show, in addition to the ROI (light gray) and the auditory activation in the right posterior parahippocampus (R pPH), the corresponding ‘visual material-processing’ results (blue) from the visual localizer experiment ($p < 0.05$, uncorrected). Mean event-related haemodynamic response functions (created by extracting percent signal change values from the R pPH) for each of the stimuli (M =material sounds, X =scrambled material sounds, H =human sounds) are plotted in the graph. Individual results from each participant are illustrated at bottom. Dashed box contains results from patient MC.

to subsequently probe the functional results of patient MC (see below).

Cortically blind participant

In light of the above results, it was also of interest to investigate how patient MC would perform on these material-processing tasks, given the widespread damage to her occipital cortex and left parahippocampus. During pre-fMRI behavioural testing, it became clear that MC was unable to perceive any visual textures. With respect to material sounds however, MC was quite adept at identifying specific materials (though she often categorized styrofoam sounds as plastic). Most importantly, MC never confused scrambled material sounds and

real material sounds. Not surprisingly, the $vM - vN$ contrast in the visual localizer fMRI experiment did not elicit any material-specific haemodynamic activity in the parahippocampal or ventro-medial occipital areas, even at a threshold criterion of $p < 0.05$, uncorrected. The story was quite different with respect to the auditory fMRI task, however, because, MC was perfect at identifying the M and H sounds, and near perfect (0.92) for the X sounds. t -tests designed to compare a single mean to a group mean (Statistica V.6, StatSoft, Inc.) confirmed that her performance was well-within the range of the participants in Experiment 1 (all comparisons $p > 0.10$). Because of her extensive lesions and lack of activation related to visual surface properties during the visual localizer experiment, an ROI could not be created to analyze the auditory data. Nevertheless, at an uncorrected threshold

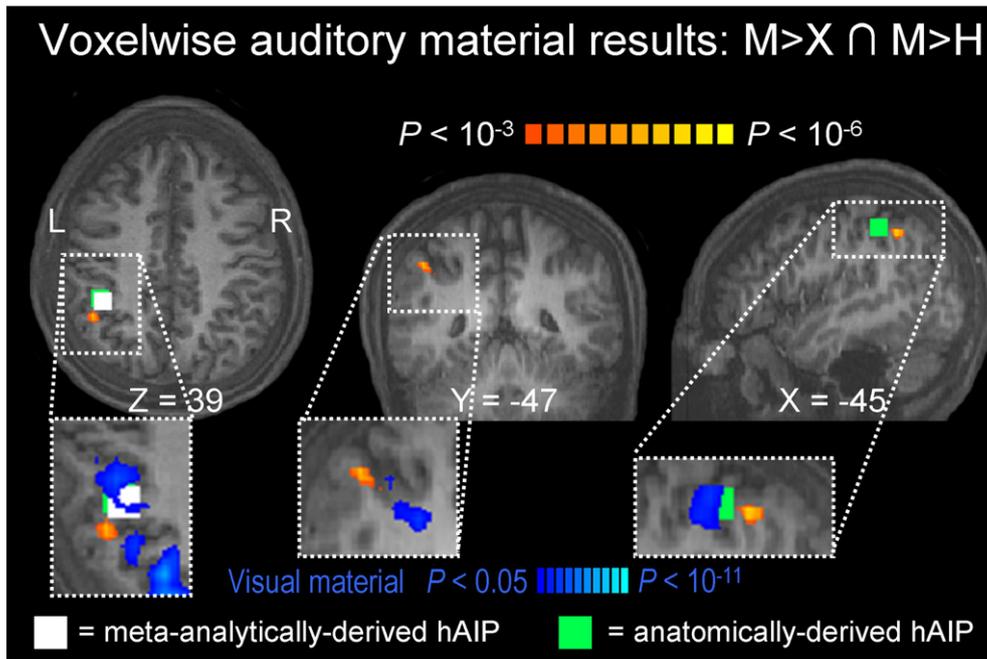


Fig. 4. Voxel-wise auditory ‘material-processing’ conjunction analysis in the neurologically intact group showing left inferior parietal lobe (L IPL) activation. White cube denotes a 10 mm³ area centred on the averaged hAIP Talairach coordinates reported in an hAIP meta-analysis (Castiello and Begliomini, 2008). Green cube denotes a 10 mm³ area centred at the mean (*N*=9) location of the anterior intraparietal sulcus and the inferior post-central sulcus junction (i.e., the putative location of the human anterior intraparietal area (hAIP)). Blue areas in the inset denote the corresponding ‘material-processing’ activation from the visual localizer experiment (*p*<0.05, uncorrected).

level of *p*<0.05, a voxel-wise analysis of the auditory material conjunction contrast (i.e., *M-X*∩*M-H*) revealed a 50 voxel area of ‘material sound’ activity in the right parahippocampal cortex (see

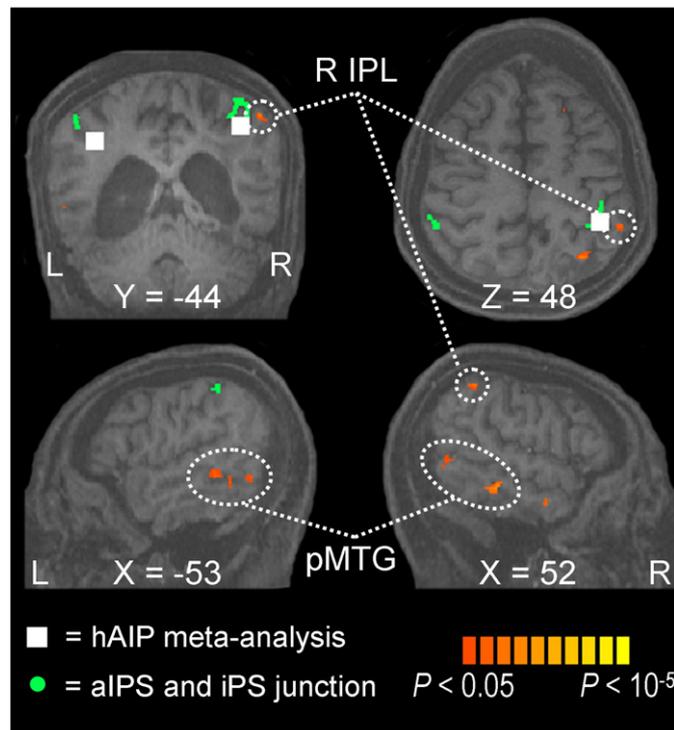


Fig. 5. Auditory ‘material-processing’ conjunction results from patient MC showing auditory activity in the right inferior parietal lobe (R IPL) and bilateral posterior middle temporal gyri (pMTG). Green regions highlight the junction of the inferior post-central sulcus (iPS) and anterior intraparietal sulcus (aIPS). White cube denotes a 10 mm³ area centred on the averaged hAIP Talairach coordinates reported in an hAIP meta-analysis (Castiello and Begliomini, 2008).

inset of Fig. 3) that was not significantly different (*t*=1.22, *p*>0.2) from the size the activations in the nine neurologically intact participants at the same threshold level. This non-difference remained (*t*=0.95, *p*>0.3) when cluster volumes were re-quantified at a less stringent threshold level of *p*=0.1. It is also worth repeating that MC’s activation cluster was based on only four runs of the experimental task, whereas all other participants had at least five and some six runs.

Unlike the neurologically intact participants, the material conjunction contrast for MC did not show any auditory material-specific IPL activation in the left hemisphere (see Fig. 5), even at threshold levels of *p*=0.20. An additional left IPL ROI analysis defined by the location of the neurologically intact group’s IPL activation also failed to uncover any material-specific auditory activation. Specifically, using each of MC’s four runs as repeated observations, the average percent signal change for the material sounds (at the second and third TRs following sound onset) were not found to be different from that of the human (*p*>0.2) or scrambled material (*p*>0.8) sounds. However, MC did show two areas of activation in the right IPL that were lateral and caudal, respectively, to the junction of the iPS and aIPS. Finally, MC also showed auditory material specific activity in the posterior middle temporal gyrus (pMTG) bilaterally, that was not observed in the neurologically intact group (see Fig. 5). A re-examination of the neurologically intact group’s individual data revealed that at the same threshold, only two of the neurologically intact participants showed similar pMTG activity in the left hemisphere, while no participant showed right pMTG activity.

Discussion

Building on findings showing that material properties of objects are processed in ventro-medial cortical areas (Cant and Goodale, 2007), the present study sought to determine if this region would also be sensitive to material properties derived from sound alone. To our knowledge, this is the first neuroimaging study to investigate the processing of material properties *per se* through sound. Relative to scrambled material sounds as well as human non-verbal vocalizations,

audio recordings of various materials being manipulated in someone's hands (i.e., paper, plastic, aluminium foil, and styrofoam) elicited greater haemodynamic activity in a small circumscribed area located on the medial side of the right parahippocampus (PH) as well as an area in the left inferior parietal lobule (IPL).

While the area within the PH had some overlap with brain areas activated when the same participants in the same experimental session passively viewed pictures of objects rendered in different materials (e.g., plastic, wood, marble, foil), it is possible that these areas are functionally distinct. For one, the visual activation never completely encompassed the area of auditory activation, even at the most liberal visual threshold settings (i.e., $p < 0.05$, uncorrected). As well, the focus of the visual material activation was located posterior to the PH, in the collateral sulcus. Indeed, at a slightly more conservative threshold (i.e., $p < 0.001$, uncorrected) there was no longer any functional overlap of the visual ROI with the auditory-elicited parahippocampal area. Further support for this argument was found when the cortically blind patient, MC, was tested on the same paradigms. Although MC showed no behavioural or fMRI evidence for the perception of material properties through vision, she was able to perceive material properties through sound. Most importantly, during the auditory fMRI task, MC showed activation in the right PH that was very similar to that observed in the group of normally-sighted individuals. Like the neurologically intact participants, MC's parahippocampal activation was situated at the same axial level as that of the group mean, on the medial aspect of the PH. Importantly, the results from MC suggest that even though pathways carrying visual material-property information to higher brain centres may be compromised, alternative routes via the auditory input can still inform these centres.

Consistent with this, the observed PH activity on the medial side of the PH corresponds closely to monkey area TH. In monkeys, the PH can be cytoarchitecturally subdivided into three sections. Proceeding in a medial to lateral brain direction these are areas TH, TFm and TFl in the macaque (Suzuki and Amaral, 2003; Lavenex et al., 2004), and areas TH, TL and TF in the rhesus monkey (Blatt et al., 2003; Blatt and Rosene, 1998). Given the close homology between entorhinal cortex in humans and rhesus monkeys (Insausti et al., 1995), it would not be surprising if the human PH, like the monkey PH, also shared a similar three-part cytoarchitectonic division. Interestingly, neuroanatomical tracer studies in the rhesus monkey show that area TH, and only area TH, has connections with auditory cortex, receiving heavy projections from anterior temporal regions (i.e., areas TS1, TS2 and TS3) as well as auditory association areas in the STG (Blatt et al., 2003). Similarly, in macaques, area TH receives inputs from auditory association areas on the convexity of the STG, and virtually no innervations from visual areas TE and TEO (Suzuki and Amaral, 1994). The area situated laterally to TH on the other hand, is distinguished by strong visual inputs from areas TE1, TE2, TEa, IPa, and FST (Blatt et al., 2003), receiving little if any auditory input. Electrophysiological recordings support this conjecture in that auditory and auditory-visual responsive neurons have been found in area TH, but not areas TL or TF, and that activity in TL and TF neurons can only be elicited by visual or somatosensory stimulation (Blatt et al., 2003). Indeed, in the present study, a similar pattern of functional response was observed, with the auditory experiment evoking material-property activation on the medial side of the PH and the visual experiment evoking material-property activity more laterally (see Fig. 3). However, given the spatial resolution of our whole-brain experimental design, it is not possible to state with certainty that our auditory and visual material regions are functionally distinct. Future fMRI investigations of multi-modal material processing would likely benefit by restricting functional data acquisition to the PH (i.e., a region-of-interest approach).

The notion that the PH plays a role in human auditory perception is not novel. Engelien and colleagues, for example, maintain that the parahippocampal gyrus, especially in the left hemisphere, plays a

prominent role in the perception of meaningful or intelligible sounds, even when those sounds are passively attended to (Engelien et al., 2000, 2006). In those particular studies, however, the contrasts were always meaningful versus non-meaningful sounds. Moreover, the 'meaningful' category included a variety of sounds such as tools, animal vocalizations, human non-verbal sounds and other auditory signals (e.g., doorbell rings). The findings from the present study on the other hand, were derived from a conjunction contrast in which meaningful sounds containing real material-property information (i.e., different materials being manipulated by a person's hands) were compared to other meaningful sounds (i.e., human non-verbal vocalizations such as sneezes, coughs, yawns, throat clears) as well as scrambled (non-meaningful) versions of the material sounds. Although not reported here, a contrast of the meaningful sounds (i.e., material sounds and human sounds) with the scrambled sounds did evoke some left PH activity. Regardless, we would note that MC, who was entirely devoid of a left PH (see Fig. 1), was still able to perform our auditory categorization task, implying that while a left PH may be activated when listening to meaningful sounds, the absence of this structure does not preclude this ability.

The PH as locus for deriving the scene gist

An obvious question that follows from this research concerns the role that the PH plays in processing material-property information. One hypothesis is that the PH makes use of surface-property information in order to build up a representation of the immediate environment. From a visual perspective, numerous behavioural studies have demonstrated that surface-based cues can be used to obtain information about the gist of a scene without the need for identifying individual objects within those scenes (Biederman et al., 1982; Moller and Hurlbert, 1996; Oliva and Schyns, 1997, 2000; Schyns and Oliva, 1994, 1997; Oliva and Torralba, 2001; Vailaya et al., 1998). This suggests that the surface properties of objects, particularly surface material-property information like texture, can facilitate scene processing. Indeed, the PH is uniquely positioned to make use of such visual information given that texture-sensitive regions are located just caudal to it, in the collateral sulcus (present results, Cant and Goodale, 2007).

Vision, however, is but one way of gathering surface-property information. As has been demonstrated in the current study, sounds convey surface-property information as well. In this context, the processing of material-specific auditory information in the PH is not surprising. Moreover, although it remains to be tested, one might also anticipate somatosensory material-property information to activate the PH. Touch-sensitive neurons are known to exist between PH areas TH and TL in the monkey (Blatt et al., 2003).

As alluded to in the introduction, one must also keep in mind that material-property information is largely derived from experience rather than the physics of the distal stimulus. In vision, the size and orientation of an object can be computed almost entirely from optics, whereas information about the intrinsic material properties of the object, such as its fragility and weight, can be predicted only on the basis of earlier experience with that object or similar objects. In other words, we have learned to associate the visual appearance of an object with haptic information about its fragility and weight, and other material properties. Similarly, whereas the location of an object can be computed from interaural timing and intensity differences, the acoustic route to the material properties depends entirely on previous associations between sound and information from other senses. Thus, an understanding of material properties comes about as a result of a rich multi-modal experience. The PH is situated just caudal to the entorhinal and perirhinal cortices, two mnemonic and memory-focused areas ideally suited for orchestrating multi-modal inputs into more abstract object representations (e.g., Buckley, 2005).

It is likely that some sort of visual imagery is invoked when people hear the sounds of objects being manipulated. However, this does not mean that one can dismiss the present results as nothing more than visual imagery. First, the area invoked by the auditory stimuli in the PH did not correspond to the focus of activation for the visual texture. Second, we observed remarkably similar auditory activation in the PH in MC despite the fact that almost her entire ventral pathway upstream from the PH was missing and she showed no differential activation for the changes in the surface properties of objects. Although one cannot rule out visual imagery in MC, it is unlikely that such imagery would be played out in early retinotopic regions in the manner proposed by [Kosslyn and Thompson \(2003\)](#).

But more importantly, as we have just discussed, the perception of material properties whether visual or auditory is entirely dependent on previous multi-modal interactions. In other words, imagery, whether visual, auditory, or somatosensory, may be part and parcel of the perception of material properties.

Why the right PH would be activated more in this task is unclear, although it has been noted that the right hippocampus is more involved in visuospatial memory as opposed to context-dependent episodic or autobiographical memory ([Burgess et al., 2002](#)). Also, the right hippocampus and PH seem to be especially important for forming an integrated representation of the extended world ([Barrash et al., 2000](#)). Whether the right PH is somehow specialized to accomplish these tasks requires further research, but it has been noted that at least for vision, there is a significant amount of overlap in the right hemisphere between the PPA and areas involved in processing material information ([Cant and Goodale, 2007](#)).

Inferior parietal lobule activation

Apart from the right PH, the only other auditory material-specific area that was uncovered in the conjunction contrast occurred in the left IPL along the IPS. A very similar type of activation has been previously reported when listeners attended to tool-related sounds compared to animal vocalizations ([Lewis et al., 2005, 2006](#)). In those studies, it was proposed that the tool sounds triggered a 'mental mimicry' of motor production sequences most likely to have produced the sounds, with the left hemispheric activation reflecting the fact that the right-handed dominant participants typically handle tools with their right hand. Further support for this hypothesis stems from the fact that the activation occurred in and around area hAIP, an area that is activated not only during real grasping movements, but during imagined movements as well as passive observation of three-dimensional objects where the viewer is not even required to pick up the object (for a review, see [Culham et al., 2006; Culham and Valyear, 2006](#)).

In the present study, the left IPL activation was observed to occur just caudal to hAIP (i.e., the IPS and aIPS junction, [Binkofski et al., 1998; Culham et al., 2003, 2006; Culham and Valyear, 2006; Faillenot et al., 1997; Frey et al., 2005](#)). Moreover, despite not ever being told how the material sounds were created (i.e., audio recordings of a person crumpling sheets of paper, plastic, aluminium foil or styrofoam in their hands), when asked to describe the sounds during debriefing, all participants (including patient MC) described something akin to a person crumpling different materials with their hands. As was the case with the previous study showing left IPL activity in response to manipulable sounds ([Lewis et al., 2005](#)), all participants were right-handed. Accordingly, the IPL activity observed in the present study may indeed be reflecting a form of 'mental mimicry' or 'action replay' that aids the listener in identifying these particular hand-produced sounds.

That said, the current study also raises the possibility that auditory-induced IPL activation may be reflecting an audio-visual-motor sequence rather than a pure audio-motor sequence. This stems from the fact that while the left IPL activity observed in the present

study was evident in the posterior IPS of the neurologically intact participants, it was noticeably absent in the cortically blind patient, MC, who also reported that the auditory material stimuli sounded like somebody crumpling different sheets of material with their hands. Bearing in mind that this single participant null finding may have been due to chance (indeed, MC showed a similar type of activation in her right IPL), it may also be the case that the absent left IPL activity was a direct result of her diminished visual capacity stemming from the virtually complete ablation of her occipital cortex some eight years earlier. The IPS, after all, is an area of the IPL known to make use of visual input, and is particularly important for integrating auditory and visual information ([Calvert et al., 2001; Macaluso et al., 2004; Meinenbrock et al., 2007](#)), especially with respect to guiding and controlling action in space (e.g., [Andersen et al., 1997; Sestieri et al., 2006](#)). As noted earlier, knowledge about auditory material properties is largely dependent on prior visual experiences, at least for normally-sighted individuals. Thus, it is plausible that the IPS auditory material-property activity reflects the integration of auditory input with its associated visual knowledge.

Interestingly, whereas patient MC did not show the IPL activity observed in the neurologically intact participants, she did show bilateral pMTG activation that was not present in the group data. Given its proximity to motion processing regions such as MT (middle temporal area) and MST (medial superior temporal area), as well as the fact that MC possesses some preserved visual motion processing ability ([Dutton, 2003](#)), the bilateral pMTG activity possibly reflects associated knowledge with motion, rather than with mental mimicry or motor plans of hand actions. Dynamic sound recordings of human footsteps have been shown to elicit activity in the posterior superior temporal sulcus (pSTS) of listeners ([Bidet-Caulet et al., 2005](#)), and both the sight or sound of manipulable tools elicit activity localized in the pMTG ([Kellenbach et al., 2003; Lewis et al., 2005; Beauchamp and Martin, 2007](#)). Along this line, it has also been shown that the right parietal lobe plays a particularly important role in auditory motion perception ([Bremmer et al., 2001; Lewis et al., 2000; Griffiths et al., 2000; Griffiths et al., 1998; Smith et al., 2007](#)), an observation that would also be consistent with MC's observed activity in the right IPL. Alternatively, right IPL activity could also reflect some sort of action replay or mental mimicry of the bimanual movements associated with handling the material, in the same way that we have speculated might be the case for the healthy participants. Clearly, while both of these ideas are in keeping with the general notion that salient properties of objects stored in memory re-activate the sensory and motor systems that were engaged when that information is initially acquired (for a review, see [Martin, 2007](#)), it must be emphasised that all of this is highly speculative. Nevertheless, further investigations could help characterize exactly what these IPL and pMTG activations represent.

In conclusion, we have demonstrated that an area of the right PH is selectively invoked when attending to the sound of object material properties. This area appears to be functionally distinct from those involved in the processing of visual material-property information, though one cannot deny that the processing of auditory and visual material-property information are intimately related to one another. At a more general level, we would speculate that the parahippocampus makes use of surface-property information in order to build up a representation of the immediate environment. In so doing, this process can stimulate action-related schema in the parietal cortex and/or the posterior middle temporal lobe.

Acknowledgments

The authors wish to thank MC for her interest and co-operation, Dr. David Purcell at the National Centre for Audiology at the University of Western Ontario for conducting audiological testing on MC, as well as two anonymous reviewers for their helpful comments.

References

- Alain, C., Arnott, S.R., Hevenor, S.J., Graham, S., Grady, C.L., 2001. "What" and "where" in the human auditory system. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12301–12306.
- Andersen, R.A., Snyder, L.H., Bradley, D.C., Xing, J., 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.* 20, 303–330.
- Arnott, S.R., Goodale, M.A., 2006. Distorting visual space with sound. *Vision Res.* 46, 1553–1558.
- Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory dual-pathway model in humans. *NeuroImage* 22, 401–408.
- Arnott, S.R., Grady, C.L., Hevenor, S.J., Graham, S., Alain, C., 2005. The functional organization of auditory working memory as revealed by fMRI. *J. Cogn. Neurosci.* 17, 1–13.
- Arnott, S.R., Heywood, C.A., Kentridge, R.W., Goodale, M.A., 2008. Voice recognition and the posterior cingulate: an fMRI study of prosopagnosia. *J. Neuropsychol.* 2, 269–286.
- Bar, M., 2004. Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629.
- Barrash, J., Damasio, H., Adolphs, R., Tranel, D., 2000. The neuroanatomical correlates of route learning impairment. *Neuropsychologia* 38, 820–836.
- Beauchamp, M.S., Martin, A., 2007. Grounding object concepts in perception and action: evidence from fMRI studies of tools. *Cortex* 43, 461–468.
- 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.
- Biederman, I., Mezzanotte, R.J., Rabinowitz, J.C., 1982. Scene perception: detecting and judging objects undergoing relational violations. *Cognit. Psychol.* 14, 143–177.
- Binkofski, F., Dohle, C., Posse, S., Stephan, K.M., Heftner, H., Seitz, R.J., Freund, H.J., 1998. Human anterior intraparietal area subserves prehension: a combined lesion and functional MRI activation study. *Neurology* 50, 1253–1259.
- Blatt, G.J., Rosene, D.L., 1998. Organization of direct hippocampal efferent projections to the cerebral cortex of the rhesus monkey: projections from CA1, subiculum, and subiculum to the temporal lobe. *J. Comp. Neurol.* 392, 92–114.
- Blatt, G.J., Pandya, D.N., Rosene, D.L., 2003. Parcellation of cortical afferents to three distinct sectors in the parahippocampal gyrus of the rhesus monkey: an anatomical and neurophysiological study. *J. Comp. Neurol.* 466, 161–179.
- Bremner, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., Fink, G.R., 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296.
- Buckley, M.J., 2005. The role of the perirhinal cortex and hippocampus in learning, memory, and perception. *Q. J. Exp. Psychol. B* 58, 246–268.
- Budinger, E., Heil, P., Hess, A., Scheich, H., 2006. Multisensory processing via early cortical stages: connections of the primary auditory cortical field with other sensory systems. *Neuroscience* 143, 1065–1083.
- Burgess, N., Maguire, E.A., O'Keefe, J., 2002. The human hippocampus and spatial and episodic memory. *Neuron* 35, 625–641.
- Calvert, G.A., Hansen, P.C., Iversen, S.D., Brammer, M.J., 2001. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage* 14, 427–438.
- Cant, J.S., Goodale, M.A., 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17, 713–731.
- Cant, J.S., Arnott, S.R., Goodale, M.A., submitted for publication. fMRI-Adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream.
- Castiello, G., Begliomini, C., 2008. The cortical control of visually guided grasping. *Neuroscientist* 14, 157–170.
- Clarke, S., Bellmann Thiran, A., Maeder, P., Adriani, M., Vernet, O., Regli, L., Cuisenaire, O., Thiran, J.P., 2002. What and where in human audition: selective deficits following focal hemispheric lesions. *Exp. Brain Res.* 147, 8–15.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Culham, J.C., Valyear, K.F., 2006. Human parietal cortex in action. *Curr. Opin. Neurobiol.* 16, 205–212.
- Culham, J.C., Dancert, S.L., DeSouza, J.F., Gati, J.S., Menon, R.S., Goodale, M.A., 2003. Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.* 153, 180–189.
- Culham, J.C., Cavina-Pratesi, C., Singhal, A., 2006. The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? *Neuropsychologia* 44, 2668–2684.
- Dutton, G.N., 2003. Cognitive vision, its disorders and differential diagnosis in adults and children: knowing where and what things are. *Eye* 17, 289–304.
- Engelien, A., Stern, E., Isenberg, N., Engelien, W., Frith, C., Silbersweig, D., 2000. The parahippocampal region and auditory-mnemonic processing. *Ann. N. Y. Acad. Sci.* 911, 477–485.
- Engelien, A., Tuscher, O., Hermans, W., Isenberg, N., Eidelberg, D., Frith, C., Stern, E., Silbersweig, D., 2006. Functional neuroanatomy of non-verbal semantic sound processing in humans. *J. Neural Transm.* 113, 599–608.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Faillenot, I., Sakata, H., Costes, N., Decety, J., Jeannerod, M., 1997. Visual working memory for shape and 3D-orientation: a PET study. *NeuroReport* 8, 859–862.
- Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.
- Fenske, M.J., Aminoff, E., Gronau, N., Bar, M., 2006. Top-down facilitation of visual object recognition: object-based and context-based contributions. *Prog. Brain Res.* 155, 3–21.
- Frey, S.H., Vinton, D., Norlund, R., Grafton, S.T., 2005. Cortical topography of human anterior intraparietal cortex active during visually guided grasping. *Cogn. Brain Res.* 23, 397–405.
- Giordano, B.L., McAdams, S., 2006. Material identification of real impact sounds: effects of size variation in steel, glass, wood, and plexiglass plates. *J. Acoust. Soc. Am.* 119, 1171–1181.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Griffiths, T.D., Rees, G., Rees, A., Green, G.G., Witton, C., Rowe, D., Buchel, C., Turner, R., Frackowiak, R.S., 1998. Right parietal cortex is involved in the perception of sound movement in humans. *Nat. Neurosci.* 1, 74–79.
- Griffiths, T.D., Green, G.G., Rees, A., Rees, G., 2000. Human brain areas involved in the analysis of auditory movement. *Hum. Brain Mapp.* 9, 72–80.
- Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7, 555–562.
- Harman, K.L., Humphrey, G.K., Goodale, M.A., 1999. Active manual control of object views facilitates visual recognition. *Curr. Biol.* 9, 1315–1318.
- Healy, E.W., Moser, D.C., Morrow-Odom, K.L., Hall, D.A., Fridriksson, J., 2007. Speech perception in MRI scanner noise by persons with aphasia. *J. Speech Lang. Hear. Res.* 50, 323–334.
- Humphrey, G.K., Khan, S.C., 1992. Recognizing novel views of three-dimensional objects. *Can. J. Psychol.* 46, 170–190.
- Insausti, R., Tunon, T., Sobreviela, T., Insausti, A.M., Gonzalo, L.M., 1995. The human entorhinal cortex: a cytoarchitectonic analysis. *J. Comp. Neurol.* 355, 171–198.
- James, T.W., Culham, J., Humphrey, G.K., Milner, A.D., Goodale, M.A., 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain* 126, 2463–2475.
- Kaas, J.H., Hackett, T.A., 1999. 'What' and 'where' processing in auditory cortex. *Nat. Neurosci.* 2, 1045–1047.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kellenbach, M.L., 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.
- Klatzky, R.L., Pai, D.K., Krotkov, E.P., 2000. Perception of material from contact sounds. *Presence* 9, 399–410.
- Kosslyn, S.M., Thompson, W.L., 2003. When is early visual cortex activated during visual mental imagery? *Psychol. Bull.* 129, 723–746.
- Lavenex, P., Suzuki, W.A., Amaral, D.G., 2002. Perirhinal and parahippocampal cortices of the macaque monkey: projections to the neocortex. *J. Comp. Neurol.* 447, 394–420.
- Lavenex, P., Suzuki, W.A., Amaral, D.G., 2004. Perirhinal and parahippocampal cortices of the macaque monkey: intrinsic projections and interconnections. *J. Comp. Neurol.* 472, 371–394.
- Lewis, J.W., Beauchamp, M.S., DeYoe, E.A., 2000. A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb. Cortex* 10, 873–888.
- Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158.
- Lewis, J.W., Phinney, R.E., Brefczynski-Lewis, J.A., DeYoe, E.A., 2006. Lefties get it "right" when hearing tool sounds. *J. Cogn. Neurosci.* 18, 1314–1330.
- Macaluso, E., George, N., Dolan, R., Spence, C., Driver, J., 2004. Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage* 21, 725–732.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U. S. A.* 92, 8135–8139.
- Martin, A., 2007. The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45.
- Meinenbrock, A., Naumer, M.J., Doehrmann, O., Singer, W., Muckli, L., 2007. Retinotopic effects during spatial audio-visual integration. *Neuropsychologia* 45, 531–539.
- Menon, V., Levitin, D.J., Smith, B.K., Lemke, A., Krasnow, B.D., Glazer, D., Glover, G.H., McAdams, S., 2002. Neural correlates of timbre change in harmonic sounds. *NeuroImage* 17, 1742–1754.
- Milner, A.D., Perrett, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heeley, D.W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E., et al., 1991. Perception and action in 'visual form agnosia'. *Brain* 114 (Pt 1B), 405–428.
- Moller, P., Hurlbert, A.C., 1996. Psychophysical evidence for fast region-based segmentation processes in motion and color. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7421–7426.
- Oliva, A., Schyns, P.G., 1997. Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognit. Psychol.* 34, 72–107.
- Oliva, A., Schyns, P.G., 2000. Diagnostic colors mediate scene recognition. *Cognit. Psychol.* 41, 176–210.
- Oliva, A., Torralba, A., 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vision* 42, 145–175.
- Peuskens, H., Claeys, K.G., Todd, J.T., Norman, J.F., Van Hecke, P., Orban, G.A., 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *J. Cogn. Neurosci.* 16, 665–682.
- Puce, A., Allison, T., Asgari, M., Gore, J.C., McCarthy, G., 1996. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J. Neurosci.* 16, 5205–5215.
- Rämä, P., Poremba, A., Sala, J.B., Yee, L., Malloy, M., Mishkin, M., Courtney, S.M., 2004. Dissociable functional cortical topographies for working memory maintenance of voice identity and location. *Cereb. Cortex* 14, 768–780.

- Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11800–11806.
- Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Schyns, P.G., Oliva, A., 1994. From blobs to boundary edges – evidence for time-scale-dependent and spatial-scale-dependent scene recognition. *Psychol. Sci.* 5, 195–200.
- Schyns, P.G., Oliva, A., 1997. Flexible, diagnosticity-driven, rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception* 26, 1027–1038.
- Sestieri, C., Di Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., Olivetti Belardinelli, M., Romani, G.L., 2006. What" versus "where" in the audiovisual domain: an fMRI study. *NeuroImage* 33, 672–680.
- Smith, K.R., Saberi, K., Hickok, G., 2007. An event-related fMRI study of auditory motion perception: No evidence for a specialized cortical system. *Brain Res.* 1150, 94–99.
- Suzuki, W.A., Amaral, D.G., 1994. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J. Comp. Neurol.* 350, 497–533.
- Suzuki, W.A., Amaral, D.G., 2003. Where are the perirhinal and parahippocampal cortices? A historical overview of the nomenclature and boundaries applied to the primate medial temporal lobe. *Neuroscience* 120, 893–906.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, New York.
- Vailaya, A., Jain, A., Zhang, H.J., 1998. On image classification: city images vs. landscapes. *Pattern Recogn.* 31, 1921–1935.
- 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., Kleinschmidt, A., Giraud, A.L., 2006. Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cereb. Cortex* 16, 1314–1322.
- Wildes, R., Richards, W., 1988. Recovering material properties from sound. In: Richards, W. (Ed.), *Natural Computation*. MIT press, Cambridge, pp. 356–363.