



## Where sound position influences sound object representations: A 7-T fMRI study

Wietske van der Zwaag<sup>a,\*</sup>, Giovanni Gentile<sup>a</sup>, Rolf Gruetter<sup>a</sup>, Lucas Spierer<sup>b</sup>, Stephanie Clarke<sup>b</sup>

<sup>a</sup> Centre for Biomedical Imaging, Université de Lausanne, Lausanne, Switzerland

<sup>b</sup> Service de Neuropsychologie et de Neuroréhabilitation, Centre Hospitalier Universitaire Vaudois and Université de Lausanne, Lausanne, Switzerland

### ARTICLE INFO

#### Article history:

Received 25 May 2010

Revised 28 September 2010

Accepted 11 October 2010

Available online 18 October 2010

#### Keywords:

Sound localisation

Non-primary auditory cortex

fMRI

Ultra-high field

### ABSTRACT

Evidence from human and non-human primate studies supports a dual-pathway model of audition, with partially segregated cortical networks for sound recognition and sound localisation, referred to as the What and Where processing streams. In normal subjects, these two networks overlap partially on the supra-temporal plane, suggesting that some early-stage auditory areas are involved in processing of either auditory feature alone or of both. Using high-resolution 7-T fMRI we have investigated the influence of positional information on sound object representations by comparing activation patterns to environmental sounds lateralised to the right or left ear. While unilaterally presented sounds induced bilateral activation, small clusters in specific non-primary auditory areas were significantly more activated by contra-laterally presented stimuli. Comparison of these data with histologically identified non-primary auditory areas suggests that the coding of sound objects within early-stage auditory areas lateral and posterior to primary auditory cortex AI is modulated by the position of the sound, while that within anterior areas is not.

© 2010 Elsevier Inc. All rights reserved.

### Introduction

Sounds produced by humans, animals or objects convey information which leads to the identification of the sound source and to its localisation in space. Sound recognition and sound localisation rely, at least partially, on distinct cortical networks, commonly referred to as the auditory What and Where (or dorsal and ventral) streams (Kaas and Hackett, 2000; Rauschecker and Tian, 2000). The separate processing of spatial and non-spatial information confers advantages, such as a very rapid discrimination of sound categories (e.g. living vs. non-living; Murray et al., 2006, or the rapidly induced, position-specific plasticity of spatial representations, Spierer et al., 2007).

How 'early' in the human sound processing stream this separation occurs is not well known. A ventral recognition stream and a dorsal localisation stream have been identified in man on the temporal and parietal convexities, respectively, in imaging (Alain et al., 2001; Anourova et al., 2001; Arnott et al., 2004; De Santis et al., 2007; Lewis et al., 2004; Maeder et al., 2001) and lesion studies (Clarke et al., 2000,

2002; Ducommun et al., 2002; Rey et al., 2007). Interactions between the auditory What and Where streams are likely to occur on the supratemporal plane, where the two streams partially overlap (Altmann et al., 2007; Arnott et al., 2004; Griffiths and Warren, 2002; Hart et al., 2004; Hunter et al., 2003; Viceic et al., 2006). Several architectonically defined areas are located within this region, including the primary (on Heschl's gyrus; Morosan et al., 2001) and several non-primary auditory areas: the anterior (AA), anterolateral (ALA), lateral (LA), medial (MA), posterior (PA), and superior temporal (STA) areas (as identified by their histo- and immunohistochemical characteristics; Rivier and Clarke, 1997; Wallace et al., 2002). The diversity of the early-stage non-primary areas may reflect functional specialisation, such as proposed for speech processing in LA and STA (Scott and Johnsrude, 2003).

In non-human primates, the separation between the two processing streams is already present within the early-stage non-primary auditory areas<sup>1</sup> of the belt; neurons within the posterior area (CL) were reported to be more selective for the location of a vocalisation stimulus than those in the anterior area (AL), which tended to be more responsive to the type of vocalisation (Tian et al., 2001). In the same study, neurons within the intermediate area (ML) were modulated by either feature, suggesting the possibility of sound

**Abbreviations:** AA, Anterior auditory area; A1, Primary auditory area; ALA, Anterolateral auditory area; BOLD, Blood Oxygenation Level Dependent; LA, Lateral auditory area; MA, Medial auditory area; MRI, Magnetic Resonance Imaging; MSP, Modulated by Sound Position; nMSP, non-Modulated by Sound Position; PA, Posterior auditory area; ROI, Region of Interest; SPM, Statistical Parametric Map; STA, Superior temporal auditory area.

\* Corresponding author. CIBM, Station 6, EPFL, CH-1015 Lausanne, Switzerland. Fax: +41 21 6937960.

E-mail addresses: [wietske.vanderzwaag@epfl.ch](mailto:wietske.vanderzwaag@epfl.ch) (W. van der Zwaag), [rolf.gruetter@epfl.ch](mailto:rolf.gruetter@epfl.ch) (R. Gruetter), [lucas.spierer@chuv.ch](mailto:lucas.spierer@chuv.ch) (L. Spierer), [Stephanie.clarke@chuv.ch](mailto:Stephanie.clarke@chuv.ch) (S. Clarke).

<sup>1</sup> The term "early-stage non-primary auditory areas" is used to refer to anatomically identified areas which, in human studies, have histological (Chiry et al., 2003; Rivier and Clarke, 1997; Wallace et al., 2002) and hodological (Tardif and Clarke, 2001) features compatible with a hierarchical level above AI but below the association cortex; they include AA, ALA, LA, MA and PA and are putative homologues of macaque belt areas.

object representation modulated by spatial information. A very early separation of the two streams is further supported by a double dissociation observed in cat. The deactivation through local cooling of the posterior or the anterior auditory fields resulted in selective deficits in sound-localisation or pattern-discrimination tasks, respectively (Lomber and Malhotra, 2008). However, to date, no non-invasive study has been able to determine or confirm the separation of What and Where streams within the auditory cortex.

The organisation of human and non-human auditory areas reveals similarities and differences. The main feature of core and belt areas is found in both (Hackett et al., 2001), while the number of belt areas appears to be different; three belt areas have been identified in macaque monkeys (Kaas and Hackett, 2000; Rauschecker and Tian, 2000), while anatomical studies identified at least five in man (Rivier and Clarke, 1997; Wallace et al., 2002). It is currently unknown to what extent human “belt” areas participate in specialised processing streams and how far their organisation is specifically adapted to speech processing.

In human subjects, several low-resolution imaging and electrophysiological studies have shown that the auditory What and Where streams overlap partially on the supratemporal plane, suggesting that the primary and early-stage non-primary auditory areas are involved in processing of either auditory feature alone or of both, potentially in combination (Altmann et al., 2007; Arnott et al., 2004; Griffiths and Warren, 2002; Hart et al., 2004; Hunter et al., 2003; Viceic et al., 2006). The spatial coding within these areas has been suggested to concern a relatively low level left/right discrimination, while the finer representation of auditory space is processed in higher-order areas, including the parietal cortex (Lewald et al., 2008).

Using high-resolution 7-T fMRI we have here investigated the influence of positional information on sound object<sup>2</sup> representations by comparing activation patterns from environmental sounds lateralised to the right or left ear. The use of high-resolution 7-T fMRI was mandatory because of the relatively small size of the auditory areas (Rivier and Clarke, 1997).

The increased signal to noise ratio, SNR, arising from the use of ultra-high magnetic field systems (>3 T), allows the use of smaller voxel sizes in functional MRI. Or, in other words, ultra-high fields facilitate image acquisition with significantly higher spatial resolution than possible at lower field strengths (Yacoub et al., 2001, 2007). The available BOLD, Blood Oxygenation Level Dependent, signal on which fMRI experiments are based, also increases with field strength, yielding even higher sensitivity (van der Zwaag et al., 2009a; Yacoub et al., 2001). Finally, the signal strength of venous blood is reduced due to a shorter relaxation time, so that spatial specificity of the BOLD signal is improved, yielding activation signals which are restricted to the cortical grey matter (Shmuel et al., 2007; van der Zwaag et al., 2009a,b). Thus, functional MRI at 7 T offers an excellent possibility to non-invasively investigate the auditory areas with high spatial precision.

## Methods

### Subjects

Eight healthy subjects (3 females, mean age  $\pm$  SD: 22  $\pm$  2 years) without history of neurological or psychiatric illness and with normal hearing provided written, informed consent to participate in the study. All procedures were approved by the Ethics Committee of the Faculty of Biology and Medicine of the University of Lausanne.

### Stimuli and paradigm

One hundred environmental sounds of living and man-made sound sources, each 1 s long, were chosen from batteries used in previous studies (Clarke et al., 1996; Murray et al., 2006); all corresponded to emotionally neutral valence and intensity (Aeschlimann et al., 2008). They were 16-bit stereo, 44,100-Hz digitisation and generated using Adobe Audition 2.0 (Adobe Systems, San Jose, CA). The sound pressure and left–right balance were controlled from the computer handling the experimental paradigm and the shielded interface unit in the magnet room. Stimuli were presented at 86–91 dB sound pressure level either to the left or to the right ear, with the same sound pressure for both ears of a volunteer (measured using a CESVA SC-L sound pressure meter; CESVA Instruments, Barcelona, Spain). Crosstalk between the left and right channels was determined to be negligible. Stimulus delivery and response recording were controlled by E-Prime (Psychology Software Tools, Pittsburgh, PA). All subjects reported satisfactory quality of the stimuli, comfortable sound levels and at least close-to-equal sound pressure in both ears.

The experimental paradigm was a triple-epoch study, consisting of right–silence–left–silence blocks. Each block lasted 30 s; those with stimuli contained two 5-s series of environmental sounds, which were presented either to the right (right) or left ear (left). During each block two image volumes were acquired. In total, 10 left- and 10 right-ear presented auditory stimuli trains were played out. The subjects were asked to attend to the stimuli during the experiment, no specific responses were required on their part. The short duration of the acquisition ensured sufficient attention levels throughout the presentation of stimuli; the post-scanning debriefing confirmed that subjects perceived the stimuli correctly.

### Hardware and scanning parameters

All subjects were scanned using an actively shielded 7 T/680 cm head-only MRI scanner (Siemens Medical Solutions, Erlangen, Germany) with an 8-channel head volume rf-coil (RAPID Biomedical GmbH, Würzburg, Germany). Auditory stimuli were delivered to the subjects using the EarPlug fMRI Hardware AudioSystem (NordicNeuroLab, Bergen, Norway). Foam padding was used to minimise scanner acoustic noise and reduce head-motion.

To minimise the effect of the scanner noise associated with the high gradient switching rate necessary for EPI acquisition on the auditory stimulation, a sparse acquisition strategy was used (Hall et al., 1999; Zaehle et al., 2007); a long repetition time allowed the auditory stimulation and the scanner acoustic noise to be well separated in time. The acquisition was timed in such a way that the superior bank of the Sylvian fissure was sampled when the maximum BOLD signal was expected, i.e. 6 s after the sound was played out. Therefore, each sound/silence instance corresponded to a single volume of the EPI acquisition train.

An EPI sequence with sinusoidal read-out (Speck et al., 2008) was used for functional data acquisition. Fifty slices aligned with the Sylvian fissure were acquired in ascending order with an in-plane resolution of 1.5 by 1.5 mm and a slice thickness of 1.5 mm. Slices were sampled in 3.14 s, inclusion of a 11.68 s delay for the sparse imaging strategy resulted in a TR of 15 s. Other parameters were: echo time/flip angle/partial Fourier factor/bandwidth = 25 ms/90°/6/8/1700 Hz/pixel in read-out direction. The phase encoding direction was anterior–posterior. Each slice was imaged with a matrix size of 128 by 128 with a field of view of 192 by 192 mm.

Anatomical images were acquired using the MP2RAGE sequence (Marques et al., 2009) with a resolution of 0.9  $\times$  0.9  $\times$  1.1 mm, a repetition time of 5 s, bandwidth 240 Hz/pixel and an echo-time of 2.87 ms. Flip angles and inversion times for the two contrasts were 4/5° and 901/3201 ms, respectively. The field of view was 230  $\times$  230  $\times$  176 mm. A

<sup>2</sup> The term “sound object” is used to refer to the acoustic experience associated with an auditory event produced by a sound source (Griffiths and Warren, 2004).

GRAPPA acceleration factor of 3 was used in the phase encoding direction. The total acquisition time was below 8 min.

### Data analysis

All fMRI datasets were preprocessed and analysed with SPM5 (Wellcome Trust Center for Neuroimaging, London, UK) for individual subjects. Pre-processing steps included motion correction, co-registration with the high-resolution anatomical MP2RAGE image and normalisation to the standard MNI template preserving the original resolution. The model was defined as boxcar functions for both conditions (left/right) and convolved with the canonical haemodynamic response. Data were high-pass filtered with a cut-off period of 240 s.

An *f*-test using the standard SPM toolbox on left- and right lateralised sounds vs. baseline in the smoothed data yielded “main effect of sounds” activation maps for each subject. For the *f*-test, images were smoothed with an isotropic Gaussian kernel with FWHM = 3 mm.

However, to retain spatial resolution and avoid the use of spatial smoothing, the model was also estimated using the wavelet-analysis SPM5 toolbox WSPM (Van De Ville et al., 2007) (subsampling scheme ‘dyadic’, transform type: ‘2D + Z’, redundancy: ‘Multiple’, wavelet flavour: ‘\*ortho’, degree (XY-plane): 1, number of iterations: 1) using the non-smoothed normalised volumes. The results were thresholded using a false discovery rate (FDR)-correction for multiple comparisons at  $p = 0.05$  (Van De Ville and Unser, 2008) and a minimum cluster size of 5 voxels.

The following SPMs were obtained for individual subjects via *t*-tests: left lateralised sounds vs. rest, right lateralised sounds vs. rest, left lateralised sounds vs. right lateralised sounds and right lateralised sounds vs. left lateralised sounds.

From these, for each hemisphere, ROIs defining areas where the activation was modulated by sound position (MSP) were formed by applying a threshold ( $p < 0.05$ , FDR corrected,  $n > 5$ ) to left lateralised sounds vs. right lateralised sounds and right lateralised sounds vs. left lateralised sounds SPMs. ROIs covering areas where signal was not modulated by position (nMSP) were formed by exclusively masking the two SPMs of lateralised sounds vs. rest with the SPMs of lateralised sounds vs. contra-lateralised sounds, using the same threshold as for the MSP ROIs.

For each condition, the location of maximum activation in MNI space and the extent of the activated area were determined using MATLAB (Mathworks Inc., Natick, MA, USA). The maxima were compared to previous literature by conversion of the map as obtained from Rivier and Clarke (1997) and Wallace et al. (2002) into MNI space via a linear regression procedure (<http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach>) and subsequent marking of the local maxima in the projection of the MNI space onto its *x*-*y* plane.

Cortical reconstruction and volumetric segmentation of the MP2RAGE anatomical images was performed with the Freesurfer image analysis suite (<http://surfer.nmr.mgh.harvard.edu/>).

### Results

An *f*-test for main effect of sounds versus baseline revealed for all individual subjects bilaterally large regions of activation on the supratemporal plane, including Heschl’s gyrus, as well as patches on the superior temporal gyrus and sulcus. Data from a representative subject are shown in Fig. 1a,b. BOLD signal changes in all subjects were essentially limited to the cortical grey matter with only few spills into the underlying white matter, as can be judged from the data in Figs. 1 and 2. This distribution of the BOLD signal, as expected from physiology, is due to the specificity of the BOLD signal in ultra-high field fMRI and the high spatial resolution employed in this study.

When comparing lateralised sound stimuli to the rest condition, both left and right-sided stimuli activated Heschl’s gyrus bilaterally as well as the surrounding parts of the *plana polare et temporale*. Sound vs. rest contrasts are shown in Fig. 1c–j as red–yellow overlays for the same representative subject as Fig. 1a–b. Activation due to right-lateralised sounds is shown overlaid on the anatomical data in Fig. 1c, e,g,i, while left-lateralised sounds are overlaid in Fig. 1d,f,h,j. This bilateral activation pattern was found for all subjects. Activated areas were, however, larger in the hemisphere contra-lateral to the presented sound. Thus, while environmental sounds compared to rest consistently presented a bilateral activation pattern, there was a general bias towards the contra-lateral hemisphere.

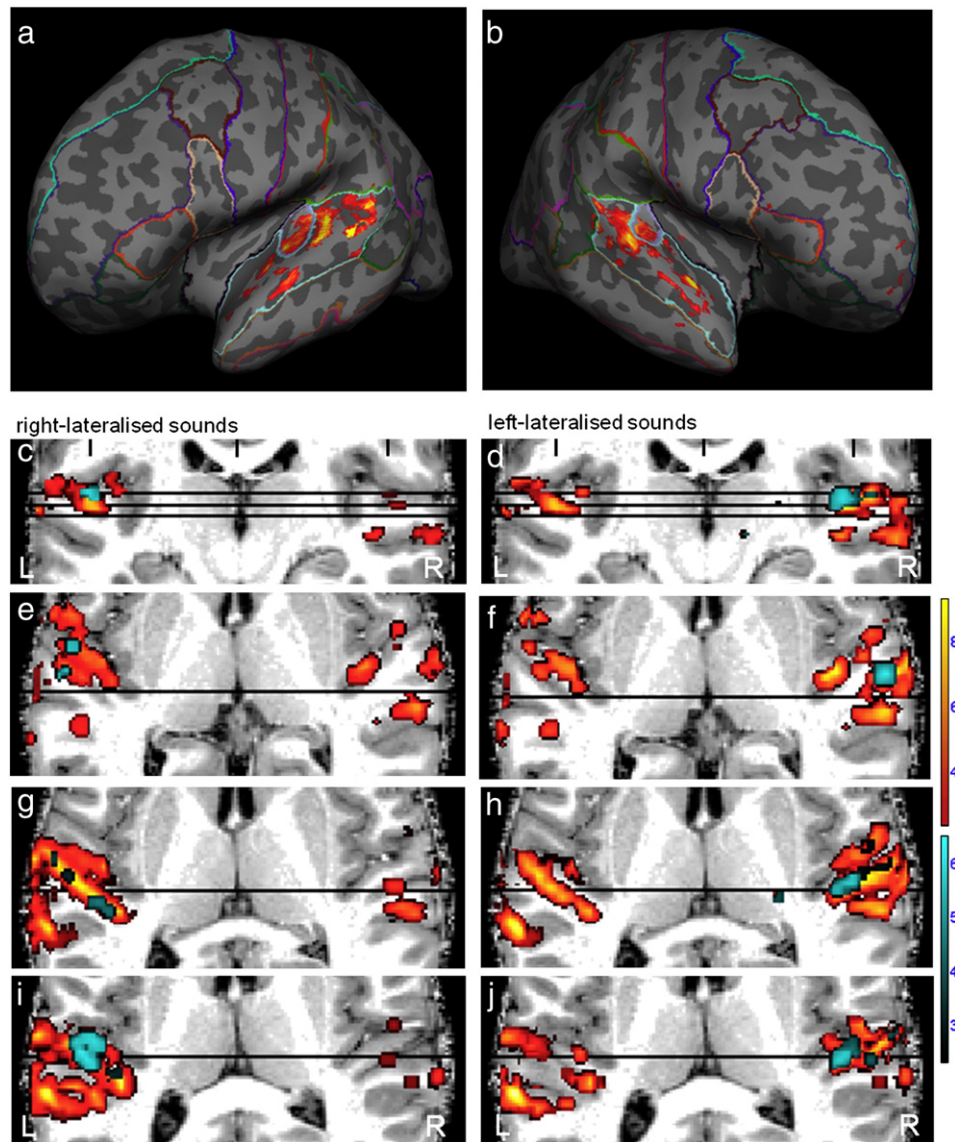
Comparison of right-lateralised vs. left-lateralised stimulus presentation revealed small clusters on the left supratemporal *planum temporale* that were significantly more activated by the right-sided stimuli (example subject, Fig. 1c,e,g,i cyan overlays). The reverse comparison, left-lateralised vs. right-lateralised stimulus presentation, revealed clusters on the right supratemporal *planum temporale* that were significantly more activated by the left-sided stimuli (Fig. 1d,f,h,j cyan overlays). These ‘modulated by sound position’ or MSP clusters were consistently detected in all 8 subjects. Examples from 4 other subjects are shown in Fig. 2. For all eight subjects, MSP regions were exclusively found with specific sensitivity for the contra-laterally presented stimulus, i.e. the right hemisphere did not yield any voxels which were significantly more activated by right-lateralised sounds than by left-lateralised sounds. MSP clusters were located in the vicinity of Heschl’s gyrus. In each hemisphere, MSP clusters covered  $1.9 \pm 0.5 \text{ cm}^2$  (mean  $\pm$  s.e.m.) of cortical surface, or  $0.9 \pm 0.2 \text{ cm}^3$  in volume.

In either hemisphere, surrounding parts of the supra-temporal plane where activation was not modulated by sound position, i.e. the nMSP areas, were much larger than the MSP clusters, reaching as much as a 20-fold volume difference (Fig. 2).

BOLD signal changes measured in each of the defined ROIs are presented in Fig. 3. A three-way repeated measure ANOVA with hemisphere (right; left); region (MSP, nMSP) and ear of presentation (left; right) as within-subject factors was conducted on the BOLD signal changes measured in the ROIs.

There was a main effect of factor region ( $F(1,7) = 41.9$ ;  $p < 0.001$ ;  $\eta^2 = 0.857$ ), indicating variation in BOLD responses as a function of ROIs types. The ANOVA further revealed interactions between factors hemisphere and region ( $F(1,7) = 5.724$ ;  $p < 0.05$ ;  $\eta^2 = 0.450$ ); hemisphere and ear ( $F(1,7) = 127$ ;  $p < 0.001$ ;  $\eta^2 = 0.948$ ) and hemisphere, region and ear ( $F(1,7) = 107.1$ ;  $p < 0.001$ ;  $\eta^2 = 0.939$ ). Follow-up two-way repeated measure ANOVAs with region and ear as within-subject factors were performed separately for the left and right hemispheres. These analyses revealed significant main effects of factors region and ear ( $p < 0.01$ ) as well significant region by ear interactions ( $p < 0.001$ ) for both the left and right hemispheres. These effects were further examined using uncorrected paired *t*-tests.

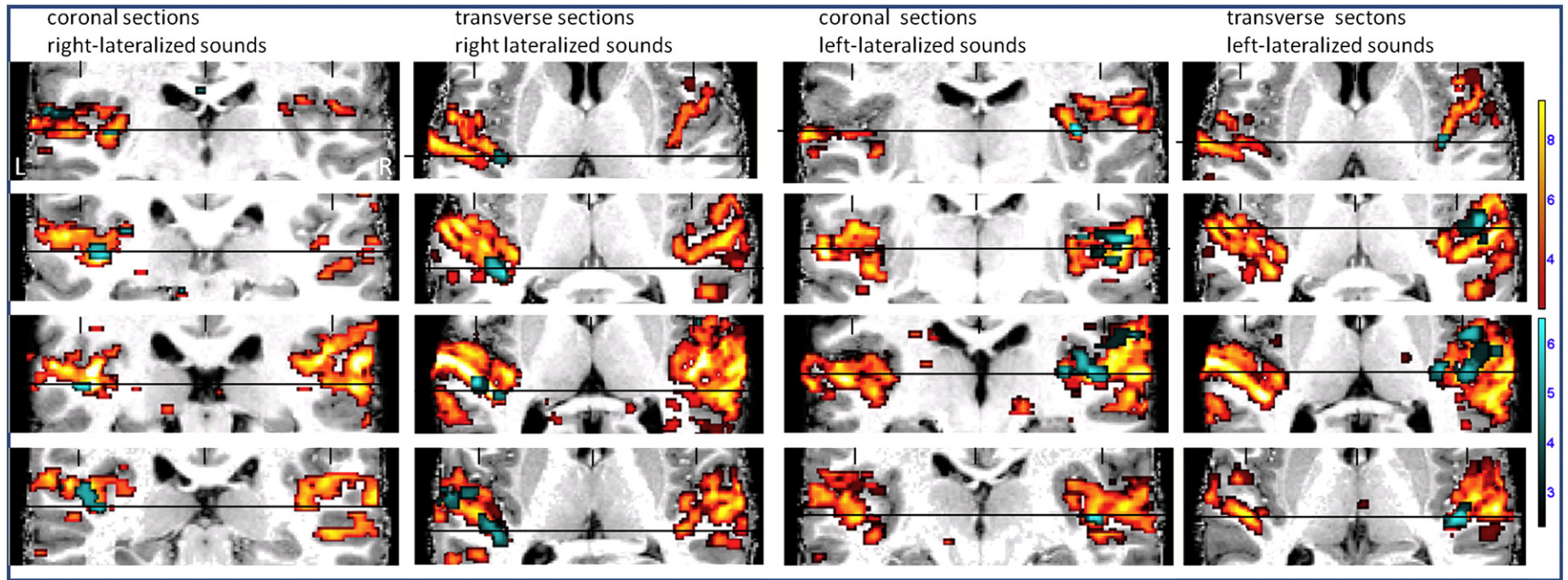
They revealed that right-lateralised stimuli yielded similar BOLD signal changes in nMSP or MSP areas in the right hemisphere ( $p > 0.001$ ), while left-lateralised stimuli resulted in highly significant differences in BOLD signal changes in right hemisphere nMSP and MSP areas ( $p < 0.0001$ ). Therefore, the differences found between signal in MSP and nMSP areas are not due to local BOLD sensitivity. Equally, for the left hemisphere, left-lateralised stimuli yielded no significant differences in BOLD signal changes in nMSP or MSP areas (paired *t*-test,  $p > 0.001$ ), while right-lateralised stimuli resulted in highly significant differences in BOLD signal in right hemisphere nMSP and MSP areas (paired *t*-test,  $p < 0.0001$ ). Interestingly, even though BOLD signal changes for all subjects were larger for right lateralised sounds in the MSP ROI in the left hemisphere than for left-lateralised sounds in the MSP ROI in the right hemisphere, this effect did not reach the  $p < 0.001$  significance level (paired *t*-test,  $p = 0.016$ ).



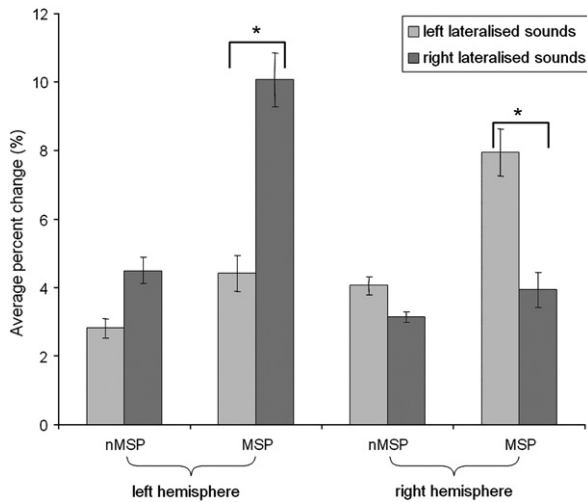
**Fig. 1.** Activation maps yielded by the presentation of environmental sounds to either ear (a, b), to the right (c, e, g, i) or left (d, f, h, j) ear. a, b: Stimulus-dependent activations; main effect of sounds in the left (a) and right (b) hemisphere of a representative subject; coloured voxels show significant activations ( $p < 0.05$ , FDR corrected) with z-scores ranging from 2.3 to 10.8 (red to yellow). Coloured lines indicate cortical parcellation segments. c–j: Effects of side of delivery (left column = right ear; right column = left ear) shown in a coronal section through Heschl's gyrus (c, d;  $y = -23$  in MNI space) and three horizontal sections (e–j);  $z = 3, 7$  and  $11$  in MNI space respectively; their positions are indicated with lines in c and d. The position of the coronal section is indicated with a line in subfigures e–j. Black ticks indicate the positions  $x = -50, 0$  and  $50$  in MNI space. Coloured voxels show significant activation to unilaterally delivered stimuli as compared to rest (red–yellow;  $p < 0.05$ , FDR corrected) or to the contralaterally delivered stimuli (cyan;  $p < 0.05$ , FDR corrected). All activation maps are shown overlaid on MP2RAGE anatomical data in MNI space. Images have been cropped to display areas of interest.

Activation map coordinates were compared to histologically identified auditory areas (Rivier and Clarke, 1997; Wallace et al., 2002). The position of local maxima from MSP activation maps of individual subjects was overlaid on a map of the supratemporal plane identifying auditory areas in both hemispheres (Rivier and Clarke, 1997) (Fig. 4). Only maxima within the auditory cortex are displayed. Within single large clusters, local maxima separated at least 8 mm were presented. Maxima of the largest cluster per subject are presented with solid labels. Overall, subjects' data presented a consistent pattern, whereby the contra- vs. ipsilateral effect was strong in areas ALA, LA and PA, with at least 5 subjects for whom a local maximum was found in the area, weak in AI and STA, with 2–4 subjects displaying a local maximum, and absent in AA and MA. For nomenclature of early-stage auditory areas, see the list of abbreviations; for the list of the MNI coordinates for each of the MSP areas of each subject, see Table 1. In comparison to the

MSP regions, larger parts of the supratemporal plane were activated by sound vs. baseline; the red outline in Fig. 4 delimits the region in which at least two subjects presented sound vs. baseline activation. Sound vs. baseline and local MSP maxima were found mostly within the architectonically identified auditory areas by Rivier and Clarke (1997) and Wallace et al. (2002) and partially outside them. A comparison with the classical cytoarchitectonic maps of von Economo and Horn (1930) indicates that sound vs. baseline activated areas TC, TD, the different subdivisions of TB as well as parts of areas TA and TG. The local maxima from MSP activations were limited to TC, TD, TB and TAB2 (which is a transition area between TB and TA). The cytoarchitectonic heterogeneity of the cortex responding to auditory stimuli outside the areas identified by Rivier and Clarke (1997) and Wallace et al. (2002) speaks in favour of other, possible higher-order auditory areas.



**Fig. 2.** Activation maps from four subjects (2, 4, 6 and 8 in Table 1), for comparison with Fig. 1. Each row shows data from a single subject. Coronal and transverse slices from maps of right-lateralised sounds versus rest (red–yellow) and right-lateralised versus left-lateralised (cyan) are shown in the two left columns, coronal and transverse slices from maps of left-lateralised sounds versus rest (red–yellow) and left-lateralised versus right-lateralised (cyan) are shown in the two right columns. Black lines indicate the position of the corresponding orthogonal slice for each activation map, black ticks indicate  $x = -50, 0$  and  $50$  in MNI space. White letters indicate left and right hemispheres. MNI coordinates of slices shown are as follows: top row, leftmost columns (lc):  $y = -23, z = 7$ , rightmost columns (rc):  $y = -17, z = 7$ . 2nd row lc:  $y = -30, z = 7$ , rc:  $y = -17, z = 7$ . 3rd row lc:  $y = -30, z = 7$ , rc:  $y = -23, z = 7$ . 4th row lc:  $y = -30, z = 7$ , rc:  $y = -23, z = 4$ .



**Fig. 3.** Activation by ipsi- and contralaterally presented stimuli (light and dark columns respectively) in MSP and nMSP regions of either hemisphere. BOLD signal changes on stimulus presentation as percentage of signal intensity at rest are shown as mean ( $\pm$  SE) of all subjects. MSP = regions modulated by sound position, i.e., with statistically significant difference between contra- vs. ipsilateral activation; nMSP = regions not modulated by sound position, i.e. regions with significant effect of sound but without a statistically significant difference between contra- vs. ipsilateral activation. The \* indicates measures which were significantly different according to paired *t*-tests ( $p < 0.0001$ ), see manuscript text.

## Discussion

Our results suggest that non-primary auditory areas on the supratemporal plane sustain multiple levels of sound object representations, of which some are independent of and others modulated by the position of the sound. A strong modulation was present within areas PA, LA and STA, located on the *planum temporale* and a weaker one within the primary auditory area and ALA; no modulation was present in areas AA and MA, located on the *planum polare* (Fig. 5). The combination of the high spatial resolution of the fMRI data, increased BOLD sensitivity and specificity available from higher field strengths (Shmuel et al., 2007; Turner et al., 1993; van der Zwaag et al., 2009a,b;

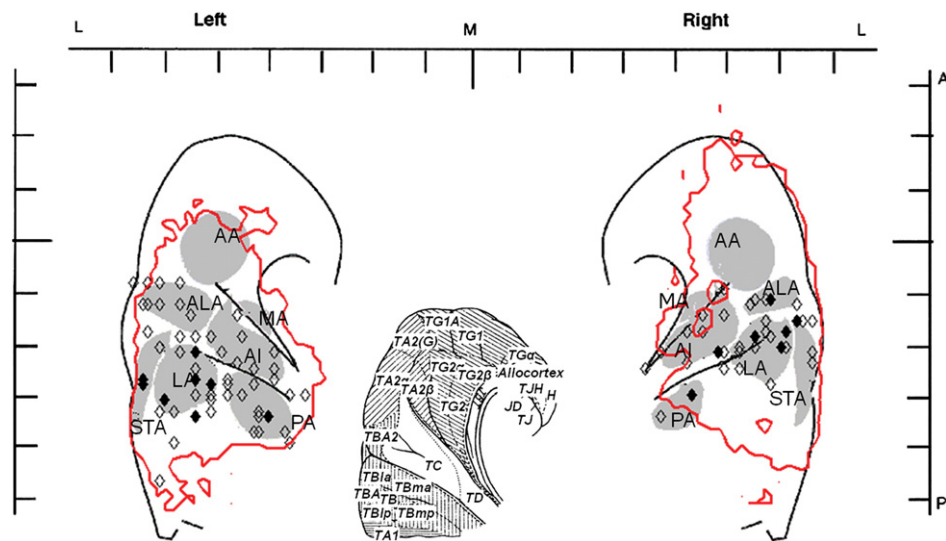
Yacoub et al., 2001) in combination with a wavelet-based analysis (Van De Ville et al., 2007; Van De Ville and Unser, 2008) of the functional images avoided a reduction of the spatial resolution. As a result, the distinction between the primary and non-primary auditory areas in fMRI data was possible and the used spatial resolution was sufficient to distinguish between areas which were sensitive to differences in presented sound location and those which were not. Found activation patterns were highly reproducible between subjects (Fig. 2).

### Early-stage auditory areas

The specific involvement of early-stage auditory areas in sound recognition and sound localisation remains poorly understood, partially due to the small size of these areas and the relatively poor spatial resolution of previous studies. Discrete parts of the *planum polare*, including areas AA and ALA (Viceic et al., 2006), were shown to be activated by tasks relevant to sound recognition or musical processes (Brown et al., 2004), but not or significantly less by tasks relevant to sound localisation (Altmann et al., 2007; Arnott et al., 2004; Griffiths and Warren, 2002; Hart et al., 2004; Hunter et al., 2003; Viceic et al., 2006). Discrete parts of the *planum temporale*, situated mainly laterally to Heschl's gyrus, (Altmann et al., 2007; Arnott et al., 2004; Griffiths and Warren, 2002; Hart et al., 2004), and more specifically areas LA, STA, and PA (Viceic et al., 2006), were shown to be involved equally in recognition and in spatial tasks (Fig. 5), as well as audio-motor coordination (Baumann et al., 2007). This agrees with our results in which spatial discrimination was especially important in LA and PA.

### Auditory responses modulated by positional information

Although spatial and non-spatial auditory information have been shown to be processed to a certain degree independently, there is also evidence for interaction. Thus, auditory spatial cues do not only contribute to auditory localisation, i.e. the ability to attribute a precise egocentric spatial co-ordinate to a sound, involving overt perception of sound location (Blauert, 1996) but play also a role in sound object segregation, sometimes referred to as 'auditory streaming' (Bregman,



**Fig. 4.** Relationship between MSP and nMSP regions and histologically identified auditory areas. Data are shown projected onto a *z*-plane in MNI space in which the auditory cortex is outlined (black solid lines) and histologically defined cortical subareas are shown in grey. For labelling of subregions, see list of abbreviations. The red solid line outlines the areas in which activation maps of at least two subjects showed significant activation for sound > rest. This area encompasses all defined subregions. Diamonds indicate local maxima from MSP activation maps of individual subjects. Solid diamonds indicate the largest MSP area per subject. Note that largest MSP clusters are located within areas PA, LA and STA; the MSP regions includes also AI and ALA, while areas AA and MA are not modulated by sound position. The inset shows the cytoarchitectonic map adapted from von Economo and Horn (1930), who subdivided the supratemporal plane into areas TC, TD, TB, TA and TG and intra- and inter cerebral cytoarchitectonic variations.



patterns associated with monoaurally delivered tones and compared them to areas identified by Rivier and Clarke (1997) and Wallace et al. (2002). Both left and right stimulations lead to bilateral activity in AI and non-primary areas; in non-primary areas PA, LA, MA, ALA and STA, analysed as a whole, a weak contralateral advantage was reported, while area AA was not analysed. The spatial resolution (3×4×5 mm) and the smoothing (5 mm at full-width half-maximum) used did not allow distinct analysis of individual areas.

By presenting environmental sounds from the right or the left side we have been able to demonstrate activation patterns that are modulated by sound position (Fig. 5). The comparison of our results with previous studies suggests that non-primary auditory areas have distinct selectivity features. Areas AA and ALA have been found to be selective for sound object recognition, as opposed to sound localisation, and were proposed to be part of the What stream (Viceic et al., 2006); see also (Altmann et al., 2007; Arnott et al., 2004; Hart et al., 2004; Hunter et al., 2003). The influence of spatial cues on responses of ALA to environmental sounds may play a role in sound object segregation; its presence within an area of the ventral (What) stream may explain the preservation of the use of spatial cues for sound object segregation in case of spatial deafness (Bellmann et al., 2001). The role of the rather strong modulation by sound position which was observed in speech related-areas LA and STA (Scott and Johnsrude, 2003) is currently not understood and needs to be established in further studies.

## Conclusions

Our results suggest that early-stage non-primary areas may be the origin of separate, specialised processing sub-streams within the classical What and Where pathways. This could be the case of two areas which belong to the What pathway and are selective for sound recognition (Viceic et al., 2006), but differ in their sensitivity to the positional information of the sound object. Area AA, which is not modulated by positional information, may carry position-independent sound object representations which lead to the ultra-rapid sound categorisation demonstrated electrophysiologically (Murray et al., 2006). Area ALA, which is modulated by positional information, may carry position-dependent sound object representations, which are relevant in situations where sound object segregation based on spatial cues or spatial attention is necessary. In this study, the positional information was mediated by ear of entry and allowed a rough left–right discrimination. Further studies are needed to determine whether precise positions, such as simulated by ITD cues or head transfer functions, also modulate sound object representations within the early stage auditory areas.

The existence of multiple levels in the representation of sound objects, here defined as position-dependent or -independent, is supported by evidence from other recent studies. Multiple levels of representation are likely to exist also for other auditory features, such as the physical properties of the sound track or stimulus familiarity, as recently indicated by fMRI priming studies (De Lucia et al., 2009). The behavioural priming effect, i.e., shorter reaction times, has been reported for words, independently of whether they were repeated by the same or a different speaker (Gagnepain et al., 2008; Orfanidou et al., 2006). The priming suppression of the BOLD response and shorter response latency occurred, however, differentially in different parts of the temporal regions, suggesting the coexistence of speaker-dependent and speaker-invariant auditory word representations. Multiple levels of object representation are not limited to the auditory modality; they have also been demonstrated in visual paradigms and in particular for view-independent and view-dependent representations within the visual What and Where pathways (James et al., 2002).

## Acknowledgments

This work was supported by the Centre d'Imagerie BioMédicale (CIBM) of the University of Lausanne (UNIL), the Swiss Federal

Institute of Technology Lausanne (EPFL), the University of Geneva (UniGe), the Centre Hospitalier Universitaire Vaudois (CHUV), the Hôpitaux Universitaires de Genève (HUG) and the Leenaards and the Jeantet Foundations and by SNF grant 3100A0-116220 to R. Gruetter and SNF grant FNS 3200030-124897 to S. Clarke.

## References

- Aeschlimann, M., Knebel, J.F., Murray, M.M., Clarke, S., 2008. Emotional pre-eminence of human vocalizations. *Brain Topogr.* 20, 239–248.
- Alain, C., Arnott, S.R., Hevenor, S., Graham, S., Grady, C.L., 2001. “What” and “where” in the human auditory system. *Proc. Natl Acad. Sci. USA* 98, 12301–12306.
- Altmann, C.F., Bledowski, C., Wibral, M., Kaiser, J., 2007. Processing of location and pattern changes of natural sounds in the human auditory cortex. *Neuroimage* 35, 1192–1200.
- Anourova, I., Nikouline, V.V., Ilmoniemi, R.J., Hotta, J., Aronen, H.J., Carlson, S., 2001. Evidence for dissociation of spatial and nonspatial auditory information processing. *Neuroimage* 14, 1268–1277.
- Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory dual-pathway model in humans. *Neuroimage* 22, 401–408.
- Baumann, S., Koeneke, S., Schmidt, C.F., Meyer, M., Lutz, K., Jancke, L., 2007. A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res.* 1161, 65–78.
- Bellmann, A., Meuli, R., Clarke, S., 2001. Two types of auditory neglect. *Brain* 124, 676–687.
- Blauert, J., 1996. *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA.
- Bregman, A.S., 1990. *Auditory Scene Analysis: The Perceptual Organisation of Sound*. MIT Press, Cambridge, MA.
- Brown, S., Martinez, M.J., Hodges, D.A., Fox, P.T., Parsons, L.M., 2004. The song system of the human brain. *Brain Res. Cogn. Brain Res.* 20, 363–375.
- Cherry, E., 1953. Some experiments on the recognition of speech, with one and two ears. *J. Acoustic Soc. Am.* 975–979.
- Chiry, O., Tardif, E., Magistretti, P.J., Clarke, S., 2003. Patterns of calcium-binding proteins support parallel and hierarchical organization of human auditory areas. *Eur. J. Neurosci.* 17, 397–410.
- Clarke, S., Bellmann, A., De Ribaupierre, F., Assal, G., 1996. Non-verbal auditory recognition in normal subjects and brain-damaged patients: evidence for parallel processing. *Neuropsychologia* 34, 587–603.
- Clarke, S., Bellmann, A., Meuli, R.A., Assal, G., Steck, A.J., 2000. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* 38, 797–807.
- 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.
- De Lucia, M., Cocchi, L., Martuzzi, R., Meuli, R.A., Clarke, S., Murray, M.M., 2009. Perceptual and semantic contributions to repetition priming of environmental sounds. *Cereb. Cortex* 20, 1676–1684.
- De Santis, L., Clarke, S., Murray, M.M., 2007. Automatic and intrinsic auditory “what” and “where” processing in humans revealed by electrical neuroimaging. *Cereb. Cortex* 17, 9–17.
- Devlin, J.T., Raley, J., Tunbridge, E., Lanary, K., Floyer-Lea, A., Narain, C., Cohen, I., Behrens, T., Zeigler, P., Matthews, P.M., Moore, D.R., 2003. Functional asymmetry for auditory processing in human primary auditory cortex. *J. Neurosci.* 23, 11516–11522.
- Ducommun, C.Y., Murray, M.M., Thut, G., Bellmann, A., Viaud-Delmon, I., Clarke, S., Michel, C.M., 2002. Segregated processing of auditory motion and auditory location: an ERP mapping study. *Neuroimage* 16, 76–88.
- Gagnepain, P., Chetelat, G., Landeau, B., Dayan, J., Eustache, F., Lebreton, K., 2008. Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *J. Neurosci.* 28, 5281–5289.
- Griffiths, T.D., Warren, J.D., 2002. The planum temporale as a computational hub. *Trends Neurosci.* 25, 348–353.
- Griffiths, T.D., Warren, J.D., 2004. What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892.
- Hackett, T.A., Preuss, T.M., Kaas, J.H., 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J. Comp. Neurol.* 441, 197–222.
- Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W., 1999. “Sparse” temporal sampling in auditory fMRI. *Hum. Brain Mapp.* 7, 213–223.
- Hart, H.C., Palmer, A.R., Hall, D.A., 2004. Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. *Hum. Brain Mapp.* 21, 178–190.
- Hunter, M.D., Smith, J.K., Taylor, N., Woods, W., Spence, S.A., Griffiths, T.D., Woodruff, P.W., 2003. Laterality effects in perceived spatial location of hallucination-like voices. *Percept. Mot. Skills* 97, 246–250.
- James, T.W., Humphrey, G.K., Gati, J.S., Menon, R.S., Goodale, M.A., 2002. Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron* 35, 793–801.
- Jancke, L., Wustenberg, T., Schulze, K., Heinze, H.J., 2002. Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation. *Hear. Res.* 170, 166–178.



- Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl Acad. Sci. USA* 97, 11793–11799.
- Lewald, J., Riederer, K.A., Lentz, T., Meister, I.G., 2008. Processing of sound location in human cortex. *Eur. J. Neurosci.* 27, 1261–1270.
- Lewis, J.W., Wightman, F.L., Brefczynski, J.A., Phinney, R.E., Binder, J.R., DeYoe, E.A., 2004. Human brain regions involved in recognizing environmental sounds. *Cereb. Cortex* 14, 1008–1021.
- Lomber, S.G., Malhotra, S., 2008. Double dissociation of 'what' and 'where' processing in auditory cortex. *Nat. Neurosci.* 11, 609–616.
- Maeder, P.P., Meuli, R.A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J.P., Pittet, A., Clarke, S., 2001. Distinct pathways involved in sound recognition and localization: a human fMRI study. *Neuroimage* 14, 802–816.
- Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2009. MP2RAGE, a self-bias field corrected sequence for improved segmentation and T(1)-mapping at high field. *Neuroimage* 49, 1271–1281.
- Middlebrooks, J.C., 2002. Auditory space processing: here, there or everywhere? *Nat. Neurosci.* 5, 824–826.
- Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13, 684–701.
- Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. *J. Neurosci.* 26, 1293–1302.
- Orfanidou, E., Marslen-Wilson, W.D., Davis, M.H., 2006. Neural response suppression predicts repetition priming of spoken words and pseudowords. *J. Cogn. Neurosci.* 18, 1237–1252.
- Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl Acad. Sci. USA* 97, 11800–11806.
- Rey, B., Frischknecht, R., Maeder, P., Clarke, S., 2007. Patterns of recovery following focal hemispheric lesions: relationship between lasting deficit and damage to specialized networks. *Restor. Neurol. Neurosci.* 25, 285–294.
- Rivier, F., Clarke, S., 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage* 6, 288–304.
- Schonwiesner, M., Krumbholz, K., Rubsamen, R., Fink, G.R., von Cramon, D.Y., 2007. Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. *Cereb. Cortex* 17, 492–499.
- Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107.
- Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K., Ugurbil, K., 2007. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *Neuroimage* 35, 539–552.
- Speck, O., Stadler, J., Zaitsev, M., 2008. High resolution single-shot EPI at 7T. *Magma* 21, 73–86.
- Spierer, L., Tardif, E., Sperdin, H., Murray, M.M., Clarke, S., 2007. Learning-induced plasticity in auditory spatial representations revealed by electrical neuroimaging. *J. Neurosci.* 27, 5474–5483.
- Spierer, L., Bellmann-Thiran, A., Maeder, P., Murray, M.M., Clarke, S., 2009. Hemispheric competence for auditory spatial representation. *Brain* 132, 1953–1966.
- Stecker, G.C., Middlebrooks, J.C., 2003. Distributed coding of sound locations in the auditory cortex. *Biol. Cybern.* 89, 341–349.
- Tardif, E., Clarke, S., 2001. Intrinsic connectivity of human auditory areas: a tracing study with Dil. *Eur. J. Neurosci.* 13, 1045–1050.
- Tardif, E., Spierer, L., Clarke, S., Murray, M.M., 2008. Interactions between auditory 'what' and 'where' pathways revealed by enhanced near-threshold discrimination of frequency and position. *Neuropsychologia* 46, 958–966.
- Thiran, A.B., Clarke, S., 2003. Preserved use of spatial cues for sound segregation in a case of spatial deafness. *Neuropsychologia* 41, 1254–1261.
- Tian, B., Reser, D., Durham, A., Kustov, A., Rauschecker, J.P., 2001. Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290–293.
- Turner, R., Jezzard, P., Wen, H., Kwong, K.K., Le Bihan, D., Zeffiro, T., Balaban, R.S., 1993. Functional mapping of the human visual cortex at 4 and 1.5 tesla using deoxygenation contrast EPI. *Magn. Reson. Med.* 29, 277–279.
- Van De Ville, D., Unser, M., 2008. False discovery rate for wavelet-based statistical parametric mapping. *IEEE J. Sel. Top. Sign. Proces.* 2, 9.
- Van De Ville, D., Seghier, M.L., Lazeyras, F., Blu, T., Unser, M., 2007. WSPM: wavelet-based statistical parametric mapping. *Neuroimage* 37, 1205–1217.
- van der Zwaag, W., Francis, S., Head, K., Peters, A., Gowland, P., Morris, P., Bowtell, R., 2009a. fMRI at 1.5, 3 and 7 T: characterising BOLD signal changes. *Neuroimage* 47, 1425–1434.
- van der Zwaag, W., Marques, J.P., Hergt, M., Gruetter, R., 2009b. Investigation of high-resolution functional magnetic resonance imaging by means of surface and array radiofrequency coils at 7 T. *Magn. Reson. Imaging* 27, 1011–1018.
- Viceic, D., Fornari, E., Thiran, J.P., Maeder, P.P., Meuli, R., Adriani, M., Clarke, S., 2006. Human auditory belt areas specialized in sound recognition: a functional magnetic resonance imaging study. *NeuroReport* 17, 1659–1662.
- von Economo, C., Horn, L., 1930. Über Windungsrelief, Maße und Rindenarchitektonik der Supratemporalfläche, ihre individuellen und ihre Seitenunterschiede. *Z. Neurol. Psychiatr.* 130, 678–757.
- Wallace, M.N., Johnston, P.W., Palmer, A.R., 2002. Histochemical identification of cortical areas in the auditory region of the human brain. *Exp. Brain Res.* 143, 499–508.
- Woldorff, M.G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., Heinz, H.J., Scheich, H., 1999. Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Hum. Brain Mapp.* 7, 49–66.
- Woods, D.L., Stecker, G.C., Rinne, T., Herron, T.J., Cate, A.D., Yund, E.W., Liao, I., Kang, X., 2009. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS ONE* 4, e5183.
- Yacoub, E., Shmuel, A., Pfeuffer, J., Van De Moortele, P.F., Adriani, G., Andersen, P., Vaughan, J.T., Merkle, H., Ugurbil, K., Hu, X., 2001. Imaging brain function in humans at 7 Tesla. *Magn. Reson. Med.* 45, 588–594.
- Yacoub, E., Shmuel, A., Logothetis, N., Ugurbil, K., 2007. Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *Neuroimage* 37, 1161–1177.
- Yost, W.A., 1991. Auditory image perception and analysis: the basis for hearing. *Hear. Res.* 56, 8–18.
- Zaehle, T., Schmidt, C.F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., Jancke, L., 2007. Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. *Neuroimage* 37, 1195–1204.