The Journal of Neuroscience

https://jneurosci.msubmit.net

JN-RM-0587-18R2

Active sound localization sharpens spatial tuning in human primary auditory cortex.

Kiki van der Heijden, Maastricht University
Josef Rauschecker, Georgetown University School of Medicine
Elia Formisano, Faculty of Psychology, Universiteit Maastricht
Giancarlo Valente, Faculty of Psychology and Neuroscience, Maastricht University
Beatrice de Gelder, Maastricht University

Commercial Interest:

This is a confidential document and must not be discussed with others, forwarded in any form, or posted on websites without the express written consent of The Journal for Neuroscience.

Title: Active sound localization sharpens spatial tuning in human primary auditory cortex.

Abbreviated title: Spatial tuning in human auditory cortex.

Authors: Kiki van der Heijden[1], Josef P Rauschecker[2], Elia Formisano[1,3], Giancarlo Valente[1], Beatrice de Gelder[1,4].

Affiliations:

1. Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD, Maastricht, The Netherlands.
2. Laboratory of Integrative Neuroscience and Cognition, Department of Neuroscience, Georgetown University Medical Center, Washington, DC 20007, USA.
3. Maastricht Center for Systems Biology, Maastricht University, 6200 MD, Maastricht, The Netherlands.
4. Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK.

Corresponding author:

Beatrice de Gelder[1], Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD, Maastricht, The Netherlands. Telephone: 0031 (0) 43 3881437. E-mail address: b.degelder@maastrichtuniversity.nl.

Number of pages: 43

Number of figures: 8

Number of tables: 1

Number of words Abstract: 220

[1] This author is also the Lead Contact for this article.

Number of words Introduction: 587

Number of words Discussion: 1,431

Conflict of interest: The authors declare no competing financial interests.

Acknowledgements:

This research was funded by a European Research Council (ERC) grant under the European Union Seventh Framework Programme for Research 2007–2013 (grant agreement number 295673), by European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 645553, ICT DANCE (IA, 2015-2017; BdG), by a PIRE Grant from the U.S. National Science Foundation (OISE-0730255; JPR), by NIH Grants R01EY018923 and R01DC014989 (JPR), and with partial support from the Technische Universität München Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763 (JPR), and NWO Vici-Grant 453-12-002 and the Dutch Province of Limburg (EF), and funding for a research exchange from the Erasmus Mundus Auditory Cognitive Neuroscience Network (KH).

ABSTRACT

Spatial hearing sensitivity in humans is dynamic and task-dependent, but the mechanisms in human auditory cortex that enable dynamic sound location encoding remain unclear. Using functional magnetic resonance imaging (fMRI), we assessed how active behavior affects encoding of sound location (azimuth) in primary auditory cortical areas and planum temporale (PT). In the current hierarchical model of auditory processing and cortical functional specialization, the planum temporale (PT) is implicated in sound location ('where') processing. Yet, strikingly, our results show that spatial tuning profiles in the left primary core and right caudo-medial belt sharpened during a sound localization ('where') task compared to a sound identification ('what') task. In contrast, spatial tuning in PT was sharp but did not vary with task performance. We further applied a population pattern decoder to the measured fMRI activity patterns, which confirmed the task-dependent effects in the left core: sound location estimates from fMRI patterns measured during active sound localization were most accurate. In PT, decoding accuracy was not modulated by task performance. These results indicate that changes of population activity in human primary auditory areas reflect the dynamic and task-dependent processing of sound location. As such, our findings suggest that the hierarchical model of auditory processing may need to be revised to include an interaction between primary and functionally specialized areas depending on behavioral requirements.

SIGNIFICANCE STATEMENT

According to a purely hierarchical view, cortical auditory processing consists of a series of analysis stages from sensory (acoustic) processing in primary auditory cortex to specialized processing in higher-order areas. Posterior-dorsal cortical auditory areas – planum temporale (PT) in humans – are considered to be functionally specialized for spatial processing. However, this model is based mostly on passive listening studies. Our results provide compelling evidence

that active behavior (sound localization) sharpens spatial selectivity in primary auditory cortex, while spatial tuning in functionally specialized areas (PT) is narrow but task-invariant. These findings suggest that the hierarchical view of cortical functional specialization needs to be extended: our data indicate that active behavior involves feedback projections from higher-order regions to primary auditory cortex.

INTRODUCTION

Sound localization is a crucial component of mammalian hearing. In the mammalian auditory cortex, neural activity in posterior areas is modulated by sound location more than in primary and anterior areas. These spatially sensitive areas include the caudo-medial and caudo-lateral belt areas (CM and CL) in non-human primates (e.g. Tian, Reser, Durham, Kustov, & Rauschecker, 2001), the posterior auditory field (PAF; Harrington, Stecker, Macpherson, & Middlebrooks, 2008) and dorsal zone (DZ) in cats (Lomber & Malhotra, 2008; Stecker, Harrington, & Middlebrooks, 2005; Stecker & Middlebrooks, 2003), and the planum temporale (PT) in humans (Brunetti et al., 2005; Deouell, Heller, Malach, D'Esposito, & Knight, 2007; Derey, Valente, de Gelder, & Formisano, 2015; McLaughlin, Higgins, & Stecker, 2016; Van der Zwaag, Gentile, Gruetter, Spierer, & Clarke, 2011; Warren & Griffiths, 2003). For this reason, cortical processing of sound location is presumably taking place in a functionally specialized, posterior-dorsal 'where' stream (Arnott, Binns, Grady, & Alain, 2004; Rauschecker & Scott, 2009; Rauschecker & Tian, 2000; Tian et al., 2001).

Behavioral evidence from psychophysical studies shows that auditory spatial sensitivity in humans is dynamic. For example, an auditory target is processed faster when auditory spatial attention is focused at the location of the target (e.g. Mondor & Zatorre, 1995; Rorden & Driver, 2001; Spence & Driver, 1994). A recent study investigating the neural mechanisms underlying this dynamic spatial sensitivity in cats identified the primary auditory cortex (A1) as a potential locus for such dynamic sound location processing. (Lee & Middlebrooks, 2011). In humans, a recent study reported a region in posterior auditory cortex that exhibited a differential level of activation based on task performance, but no task modulation of selectivity to interaural level (ILD) or time differences (ITD) across the entire auditory cortex (Higgins, McLaughlin, Rinne, & Stecker, 2017). However, it is presently not clear whether task performance results in sharpening of spatial tuning within distinct regions of the human auditory cortex, and whether

this sharpening occurs preferentially in functionally specialized ‘where’ regions (i.e. PT) or also affects primary auditory cortex.

Moreover, the effects of task performance on the cortical encoding of sound location are not yet known. The computational mechanisms underlying cortical sound location encoding are still a matter of debate, and prior studies assessing the validity of these computational mechanisms have not addressed possible effects of task performance (Day & Delgutte, 2013; Derey et al., 2015; Harper & McAlpine, 2004; King et al., 2007; McAlpine, Jiang, & Palmer, 2001; Miller & Recanzone, 2009; Ortiz-Rios et al., 2017; Stecker et al., 2005; Stecker & Middlebrooks, 2003).

Here we measured with fMRI the neuronal population responses to different sound azimuth positions in the human auditory core, lateral belt areas, and planum temporale (PT), while participants performed different behavioral tasks. We then evaluated the spatial selectivity of neuronal populations within these areas across task conditions. Additionally, we applied a modified version of a maximum-likelihood population-pattern decoder previously used to decode sound location from neural spike rates (Day & Delgutte, 2013; Jazayeri & Movshon, 2006; Miller & Recanzone, 2009) to assess whether sound location encoding in fMRI activity patterns in human auditory cortex within and across hemispheres is modulated by task performance. Our results provide new insights into the dynamic nature of sound location encoding in primary human auditory cortex. In particular, in agreement with “reverse hierarchy” (Ahissar, Nahum, Nelken, & Hochstein, 2009) and “recurrent processing” models (Bullier, 2001; Lamme & Roelfsema, 2000), our data suggest that behavior (sound localization) is enabled by feedback from functionally specialized areas to primary auditory cortex.

MATERIALS AND METHODS

Participants

Thirteen human volunteers gave informed consent to participate in the experiment. Data of two participants were excluded from the analysis due to insufficient data quality as a consequence of excessive motion and participant fatigue. Data of the remaining eleven participants (mean age = 28.9 years, standard deviation = 11.7 year, seven females) are presented here. Participants reported no history of neurological disorders. We assessed hearing levels with pure-tone thresholds (0.5, 1, 2, 4, 8 kHz) using an Oscilla SM910 Screening Audiometer (Oscilla, Aarhus, Denmark). Hearing thresholds did not exceed 25dB for any of the frequencies tested. The Institutional Review Board (IRB) of Georgetown University granted approval for the study.

Stimuli

Stimuli consisted of amplitude-modulated (AM) white noise clips (*probe* sounds, duration = 1200 ms) and click trains (*target* sounds, click rate = 200 Hz, duration = 1200 ms). Probe and target sounds were created with Matlab (The Mathworks, Inc., Natick, Massachusetts, United States). Stimuli were presented at one of seven locations (-90°, -60°, -30°, 0°, +30°, 60°, and +90°; Figure 1 A.

All stimuli were spatialized by making subject-specific binaural recordings (Derey et al., 2015). During the binaural-recording session, participants sat in a chair in the center of a production studio (internal volume = 66 m^3, walls and ceiling consisted of gypsum board covered with fabric, the floor consisted of concrete covered with a carpet) with binaural microphones placed in their ear canals (OKM II Classic Microphone, Soundman, Germany). A loudspeaker positioned at zero elevation in the far field (distance to subject = 1.3 meters) presented sounds

at each of the locations (Figure 1 A). This procedure resulted in stimuli with a clear spatial percept based on available ILD, ITD, and spectral cues (Figure 1 C and Figure 1 D).

Each stimulus was pre-filtered with headphone equalization filters provided by the manufacturer of the MRI-compatible earbuds used in the present study (Sensimetrics S14; Sensimetrics Corporation, Cambridge, MA, United States). The headphone equalization filters ensure a flat frequency response at the level of the earbuds and remove headphone-induced phase offsets between the earbuds.

For the tonotopy measurements, we used amplitude-modulated pure tones (rate of modulation = 10 Hz, full-depth modulation, 800ms duration). Pure tones were centered on eight center frequencies (0.18; 0.30; 0.51; 0.86; 1.46; 2.48; 4.19; 7.09 kHz) with a slight variation of ±0.1 octave to prevent habituation (see De Martino et al., 2013). Stimuli for the tonotopy measurements were also pre-filtered with the headphone equalization filters described above.

Experimental design

Participants listened to *probe* trials in three behavioral conditions: *passive listening, sound identification*, and *sound localization. Probe* trials consisted of five repetitions of a *probe* sound clip (duration = 1200ms) at the same location. Sound clips were presented in silent gaps (1.4s) in between fMRI acquisition periods (2s, see *Data acquisition*), resulting in a total duration of 17 seconds per trial (five stimulus repetitions in silent gaps of 1.4s plus five fMRI data acquisition periods of 2s; Figure 1 B). In the active listening conditions only, participants also listened to *target* trials. Specifically, in the *sound identification* condition, target trials had a similar structure (i.e. five repetitions at the same azimuthal location), yet the fourth or the fifth repetition of the *probe* sounds (AM white nosie) was replaced by a deviant *target sound* (click train) at the same location (Figure 1 B). In the *sound localization* condition, target trials had a similar structure as well, but the fourth or the fifth repetition of the *probe* sound (AM white noise) was replaced by a

probe sound at a deviant azimuth location. For example, the first four stimuli were presented at -90º and the fifth stimulus at +30º (Figure 1 B).

During fMRI acquisition, trials were grouped by task (passive listening, sound identification, sound localization) in a block. In each block, probe trials were presented once at each azimuth location and were separated by an inter-trial interval of 12.2s (see *Data acquisition* for detailed information). The order of azimuth locations was randomized within a block. Thus, for *passive listening*, a block consisted of seven probe trials, one at each azimuth location. For the active tasks – *sound localization* and *sound identification* – a block also contained two target trials (equivalent to ~22% of the total number of trials) in addition to the seven probe trials. The order of target and probe trials was randomized within a block.

Each participant performed one block of each task per run of fMRI acquisition. Thus, one run consisted of three blocks corresponding to the three behavioral task conditions. At the start of each task block, a short audio clip of a voice informed participants of the task at hand: 'sound location', 'sound identity', or 'passive listening'. In the *passive listening* condition, participants listened to the sounds without making a response. In the *sound identification* condition, participants pressed a button immediately upon detection of a target sound within a target trial (i.e. the click train). In the *sound localization* condition, participants pressed a button immediately upon detecting a location switch within a target trial.

The order of blocks was randomized and counterbalanced across participants. In total, participants completed four runs of the main experiment (~10 minutes each) in the MRI scanner. This resulted in four probe trial repetitions per azimuth location per task condition. Only probe trials were included in the subsequent analyses (see *Data analysis*).

Prior to the fMRI measurements, participants performed a short practice session to get familiar with the tasks and with the MRI environment. This also enabled participants to get accustomed

to the auditory spatial percept in a supine frame of reference (due to the supine position required by the MRI scanner). The practice session consisted of passive presentation of the probe stimuli at each location as well as short task blocks of the *sound localization* and the *sound identification* task, in which one target trial was presented per task block.

Finally, the scan session was concluded with two runs of tonotopy measurements (~7.5 minutes each). For this experiment, participants listened passively to blocks of AM pure tones in the MRI scanner. Each block was repeated twice per run, resulting in four repetitions per center frequency. The order of frequency blocks was randomized (see De Martino et al., 2013).

Data acquisition

Data were acquired with a Siemens TIM Trio 3-Tesla magnetic resonance imaging (MRI) scanner at the Center for Functional and Molecular Imaging (CFMI) at Georgetown University (Washington, DC, United States). For the main experiment, blood-oxygenated-level-dependent (BOLD) signals were measured with a T_2*-weighted echo-planar imaging (EPI) sequence covering the temporal cortex and parts of the occipital, parietal, and frontal cortex (echo time [TE] = 30 ms; repetition time [TR] = 3400 ms; flip angle = 90°; number of slices = 32; voxel size = 2 mm^3 isotropic). Image acquisition was clustered (TA = 2000 ms), and binaural recordings were presented in silent gaps (duration = 1400 ms) between subsequent volume acquisitions through MR-compatible insert earphones (Sensimetrics S14, Sensimetrics Corporation) with sound-attenuating foam eartips (>29dB attenuation). One sound was presented per TR. Trials (i.e. five stimulus repetitions per azimuth location corresponding to five TRs, 17 s duration) were separated by three volumes in which no sound was presented (that is, 12.2 s silence) to allow the BOLD signal to return to baseline before the onset of the next trial.

We also acquired a high resolution anatomical image of the whole brain with a MPRAGE T1-weighted sequence (TE = 2.13 ms; TR = 2400 ms; voxel size = 1 mm^3 isotropic). For the

tonotopic measurements we also used a sparse T_2*-weighted EPI sequence to measure the BOLD signal, covering mainly the temporal cortex (echo time [TE] = 30 ms; repetition time [TR] = 2600 ms; acquisition time (TA) = 1600 ms; silent gap = 1000 ms; flip angle = 90°; number of slices = 25; voxel size = 2 mm^3 isotropic). In each run, AM pure tones were presented in the silent intervals between subsequent volume acquisitions in blocks of six repetitions per center frequency (15.6 s). Blocks were separated by 12 s of silence (four volumes).

Statistical Analysis

Data preprocessing

Functional and anatomical data were analyzed using BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands), and customized Matlab code. Preprocessing of functional images included motion correction (trilinear/sinc interpolation, we used the first run of first volume as reference volume for aligning), slice scan time correction (sinc interpolation), linear drifts removal, temporal high pass filtering (threshold = 7 cycles per run), and mild spatial smoothing (3 mm kernel). Functional images were co-registered to the anatomical T1-weighted image and transformed to 3D Talairach space (Tournoux & Talairach, 1988). Gray-white matter boundaries were defined with the BrainVoyager QX automatic segmentation procedure and manually improved when necessary.

Group analyses were performed in surface space to ensure optimal alignment of the auditory cortex across participants. To this end, we applied cortex based alignment (CBA) to the surface reconstruction of each participant (Goebel, Esposito, & Formisano, 2006) with the additional constraint of an anatomical definition of Heschl's gyrus (HG; Kim et al., 2000; Morosan et al., 2001). High-resolution surface mesh time courses were created by sampling and averaging for each point on the surface (that is, each vertex) the values from -1 mm below the gray/white matter boundary up to 2 mm in the gray matter towards the pial surface.

233 Univariate analysis of the processing of spatialized sounds

234 To test for the general response to presentation of spatialized sounds, we estimated a random
235 effects general linear model (RFX GLM) with a predictor for sound presentation including all
236 probe trials (irrespective of azimuth location or behavioral task condition). Target trials were
237 modeled with a separate predictor and not included in the contrast.

238 Response azimuth functions

239 We constructed a response azimuth function (RAF) for each auditory responsive voxel
240 (individual subject GLM with one predictor per sound azimuth location per task condition and
241 excluding target trials, contrast *auditory stimuli* > *baseline*, q[FDR] < 0.05]). RAFs consisted of
242 location-specific beta values estimated with a GLM with one predictor per sound location per
243 task. RAFs were mildly smoothed with a moving average window of three points (weights [.2 .6
244 .2]). A peak response was defined as a response at 75% or more of the maximum beta value in
245 the RAF (see also Derey et al., 2015; Stecker et al., 2005; Stecker & Middlebrooks, 2003). Each
246 peak was described as a vector with *length* = β, and *angle* = *azimuth position*. The vector sum
247 then consisted of the summation of these individual vectors.

248 We considered a voxel to be spatially selective if the BOLD response was modulated by sound
249 azimuth position – as reflected in the RAF – such that at least one and maximally three adjacent
250 azimuth positions elicited a peak response. A voxel that exhibited a peak response to more than
251 three adjacent azimuth positions was considered omni-responsive and therefore nonselective.
252 Voxels that exhibited a peak response to two or more separate azimuth locations were also
253 considered nonselective.

254 The tuning width of spatially selective voxels was quantified as the equivalent rectangular
255 receptive field (ERRF) width (Lee & Middlebrooks, 2011). The ERRF is equal to the ratio
256 between the amplitude of the peak response (that is, the beta value at the preferred location),

and the integral of the RAF. Although this measure does not provide an absolute measure of spatial selectivity, it enables the comparison of spatial selectivity across conditions, areas and participants. Given that the rostral belt areas were not extensively activated, we focused this analysis on the caudal belt areas CM and CL.

Response sharpening versus response gain

We tested whether sharpening of spatial tuning resulted from *BOLD response gain* (that is, an increase of the BOLD response at the voxel's preferred location), *BOLD response sharpening* (a decrease in the BOLD response at the voxel's least preferred location), or a combination of the two. For this comparison, we defined the voxel's best location as the location with the highest beta value in the task-independent RAF, that is, the average RAF across the two active task conditions. Similarly, we considered the least-preferred location the azimuth location with the lowest beta value in the average RAF (see also Lee & Middlebrooks, 2011, 2013)).

Decoding sound azimuth position from fMRI activity patterns

To decode sound location, we applied a population-pattern decoder to the measured fMRI activity patterns in two regions of interest: the core region and PT. We selected these regions based on prior research in animals indicating primary auditory cortex as a potential locus for dynamic spatial sensitivity (Lee & Middlebrooks, 2011) and prior neuroimaging research in humans illustrating the role of PT in spatial auditory processing in the human brain (Brunetti et al., 2005; Deouell, Heller, Malach, D'Esposito, & Knight, 2007; Derey, Valente, de Gelder, & Formisano, 2015; McLaughlin, Higgins, & Stecker, 2016; Van der Zwaag, Gentile, Gruetter, Spierer, & Clarke, 2011; Warren & Griffiths, 2003).

In general, the decoder – a modified version of a pattern decoder introduced to decode sensory information from neural spike rate patterns (Jazayeri & Movshon, 2006; see also Day & Delgutte, 2013; Miller & Recanzone, 2009) – computes the log-likelihood that a sound at a given

azimuth location elicited the observed fMRI activity pattern. In particular, we computed for each voxel the log-likelihood that a stimulus at a particular azimuth location induced the observed BOLD response. The population log-likelihood then consists of the sum of the log-likelihoods across all voxels (Figure 2).

Specifically, for each cortical area, we selected those voxels that responded to sounds (GLM *sound* > *baseline*, $p < 0.005$ uncorrected) and exhibited a spatially selective response (see section before). Next, we estimated for each subject a GLM per functional data run with one predictor per azimuth position per task. This resulted in four beta estimates per azimuth position, equivalent to the four functional runs. Beta estimates were normalized between 0 and 1 across the seven azimuth positions within each run. For each stimulus azimuth position, we then computed the log-likelihood that the observed BOLD response (β_i) in the voxel under consideration was elicited by the presentation of a sound at that location. Assuming that the observed BOLD response β_i of voxel i for a given azimuth position θ_0 is normally distributed with mean $\mu_{0,i}$ and standard deviation $\sigma_{0,i}$, the log-likelihood of the observation can be computed as:

$$log\ L_i(\theta_0) = \frac{\left(\beta_i - \mu_{0,i}\right)^2}{2\sigma_{0,i}^2} - \frac{1}{2}\log\left(\sigma_{0,i}^2\right) - \frac{1}{2}\log(2\pi)$$

The estimation was carried out using cross-validation: we considered three runs to estimate the mean $\mu_{0,i}$ and standard deviation $\sigma_{0,i}$ of a given voxel and azimuth position, and we used the left-out run to calculate the log-likelihood. The procedure was repeated for all the possible train-test combinations. Due to the limited amount of available data (one trial per run), the estimation of the parameters was done using the beta values of the selected voxel, as well as the six neighboring voxels, that is, those voxels sharing a side with the relevant voxel. Consequently, the number of data points to estimate $\mu_{0,i}$ and $\sigma_{0,i}$ was 21 (three functional runs multiplied with seven voxels). The test data β_i is the beta estimate for this voxel for this azimuth position in the

run that was left out. Assuming conditional (i.e. within each azimuth position) independence between different voxels, the population response was then computed as the sum of log likelihood of all voxels in the cortical area (N):

$$\log L(\theta_0) = \sum_{i=1}^{N} \log L_i(\theta_0)$$

In the test run, we predicted the sound azimuth location of a new, unseen sound, by selecting the location with the highest log-likelihood. This is equivalent to using a probabilistic classifier based on the posterior probability of azimuth location given the observed data, when class prior is uniform across all sound locations. Reported absolute errors are the average across the four train-test estimations. Statistical comparisons of absolute error across cortical areas and tasks were made with Wilcoxon signed-rank tests (one-tailed) and corrected for multiple comparisons with the false discovery rate (q[FDR] < 0.05) unless mentioned differently. We determined the chance level of absolute error per azimuth position with permutation testing. Specifically, within each run we permuted beta estimates randomly across the seven azimuth locations and for all voxels independently. We then applied the maximum likelihood decoder to the permuted data. This procedure was repeated 1500 times per subject. Chance level of absolute error was computed as the mean absolute error across permutations.

Finally, we applied the population pattern decoder to data from both hemispheres simultaneously. In particular, we randomly sampled half of the voxels in the left hemisphere and half of the voxels in the right hemisphere. This procedure ensured that the number of data points used for the maximum likelihood estimation was equal when the decoder operated on data from two hemispheres versus data from a single hemisphere. We repeated the random sampling procedure 200 times per subject and computed absolute error as the average across

samples. To determine the chance level for the population decoder operating on data from the two hemispheres, we applied a similar permutation procedure as described above. However, due to the interaction of the computationally intensive procedure of repeating the random sampling of half of the voxels in each hemisphere as well as the permutations, we limited the calculation to 30 random samples with 10 permutations each. Chance level of absolute error was computed as the average absolute error across samples and permutations.

Parcellation of the auditory cortex

To divide the auditory cortex into core, belt regions, and PT, we combined maps of frequency preference (tonotopy) and frequency selectivity. To construct these maps, we first estimated a voxel's frequency tuning profile by estimating GLM with one predictor per center frequency for each auditory active voxel (assessed with a GLM contrasting *auditory stimulation* > *baseline,* liberal threshold of $p < 0.05$ uncorrected). We inferred a voxel's preferred frequency (PF) from the frequency tuning profile. That is, a voxel's PF was defined as the frequency with the highest beta value in the tuning profile (after z-normalizing across voxels). We then created tonotopic maps on the cortical surface by color coding the PF of all auditory responsive voxels in a blue (high frequency) to red (low frequency) color scale.

Next we estimated the frequency selectivity of a voxel by computing a *Frequency Selectivity Index* (FSI). This index expresses the ratio between the peak beta value (that is, the beta corresponding to the PF) and the area under the frequency tuning curve (the integral):

$$FSI = \frac{\int frequency\ tuning\ curve}{\beta_{PF}}$$

Then – similar to (Moerel, De Martino, & Formisano, 2012) – we defined the tuning width (TW) of a voxel as:

$$TW = PF / f_2 - f_1$$

where (f_2-f_1) is the frequency range in Hz corresponding to the FSI. As such, TW is high for voxels with a narrow tuning profile and small for voxels with a broad tuning profile. We color coded TW on the cortical sheet in a yellow (broad tuning) to purple (narrow tuning) color scale.

Finally, we used these maps to parcellate the auditory cortex following criteria based on the tonotopic organization described by Moerel et al. (2012); Figure 3). Specifically, Moerel et al. (2012) identify the core region as a region overlapping with HG that is narrowly tuned to frequency and encompasses two mirror-symmetric tonotopic gradients (see also Formisano et al., 2003; Leaver & Rauschecker, 2016; Moerel, De Martino, & Formisano, 2014). This core region is flanked by broadly tuned regions both anteriorly (overlapping with the first transverse sulcus and planum polare in general), and posteriorly (coinciding with Heschl's sulcus [HS]). Here we defined these broadly tuned bands as the rostral and caudal belt respectively (Figure 3). We then evenly divided both the caudal and the rostral belt into medial and lateral parts, resulting in four belt areas: caudomedial (CM), caudolateral (CL), rostromedial (RM), and rostrolateral (RL; Kaas & Hackett, 2000; Rauschecker, Tian, & Hauser, 1995). Finally, in line with Moerel et al. (2012) and the anatomical definition of PT provided by Kim et al. (2000), we defined the remaining posterior part of the superior temporal plane as PT. This region was bordered anteriorly by the caudal belt (overlapping largely with HS), medially by the insular cortex, and laterally by the superior temporal gyrus (STG).

Note that two participants did not show extensive activation in the auditory cortex for the contrast *auditory stimulation > baseline* as a result of excessive movement during the tonotopy measurements (possibly due to participant fatigue). We parcellated the auditory cortex of these two participants based on anatomical criteria, resulting in areas that were similar in size and location to those of the other participants. Specifically, the core region was identified as approximately two-thirds of HG (starting from the medial border; Moerel et al., 2012, 2014). The caudal belt was defined by HS, bordered posteriorly by PT (Kim et al., 2000). The rostral belt

was defined as anteriorly to HG – mainly overlapping with the first transverse sulcus – as the mirror image of the caudal belt. The rostral and caudal belt regions were evenly split into a lateral and medial part.

Maps of cortical auditory areas constructed in surface space were projected back into volume space. In subsequent analyses, we included for each area the voxels that responded to sounds (established with a GLM, contrast *auditory stimulation* > *baseline*, liberal threshold of $p < 0.005$ uncorrected; see Table S1).

RESULTS

Behavioral task performance

Behavioral accuracy in the MRI scanner was high for both active tasks. The average hit rate for the sound localization task was 94.3% (standard deviation [SD]: 15.2%), and for the sound identification task 90.9% (SD: 12.6%). There was no difference in mean accuracy between tasks (paired samples t-test, $t(10) = 0.607$, $p = 0.557$).

Univariate analysis of the processing of spatialized sounds in human auditory cortex

A random effects general linear model (RFX GLM) contrasting *auditory stimulation* > *baseline* showed increases in BOLD signal in primary and secondary auditory cortices in response to the probe trials (corrected for multiple comparisons with the False Discovery Rate [FDR], $q < 0.05$; (Benjamini & Hochberg, 1995). Activated areas included Heschl's gyrus (HG), HS, planum temporale (PT), and – to a lesser extent – the first transverse sulcus and other parts of the planum polare (PP). To investigate differences in the overall level of activation elicited by the three task conditions, we computed several balanced contrast maps. However, none of these contrasts revealed different activation levels between task conditions, either at a stringent

threshold (FDR, $q < 0.05$) or at a more liberal threshold ($p < 0.005$ uncorrected), indicating that the overall BOLD signal amplitude in the auditory cortex was similar across tasks.

Parcellating the human auditory cortex

In agreement with prior tonotopic mapping studies (e.g. Da Costa et al., 2011; Formisano et al., 2003; Leaver & Rauschecker, 2016; Moerel et al., 2012; Striem-Amit, Hertz, & Amedi, 2011; Talavage et al., 2004; Wessinger et al., 2001), cortical maps of frequency preference revealed a region tuned to low frequencies overlapping partly with HG which was bordered anterolaterally and posteromedially by regions responding maximally to high frequencies (Figure 3). Further, similar to Moerel et al. (2012) we observed a narrowly tuned region overlapping with (or in close vicinity to) HG in the frequency selectivity maps of most participants. This region was flanked by areas with broad frequency selectivity profiles (Figure 3). We combined these maps of frequency preference and selectivity and derived an operational definition of the core region, the belt regions (see Rauschecker et al., 1995) for original definitions in macaque auditory cortex), and planum temporale (PT; Figure 3; Table 1; see also *Methods*).

Spatial selectivity in human auditory cortex is higher in posterior, higher-order regions than in primary regions

To start, we examined general differences in the presence of spatially selective voxels between cortical areas, i.e. inter-area differences irrespective of behavioral demands. The results show that the average proportion of auditory responsive voxels that was spatially selective (averaged across task conditions) varied across cortical regions in the left hemisphere (Figure 4 A, left panel), as well as in the right hemisphere (Figure 4 A, right panel). In particular, in the left hemisphere, PT contained relatively more spatially selective voxels than the core, CM, and CL (Table 3). The proportion of selective voxels was also higher in left CL than in the left core. In

the right hemisphere, PT contained a higher proportion of selective voxels than the core and CL as well, and the proportion of spatially selective voxels was higher in CM than in CL.

We also assessed spatial selectivity by investigating the relative tuning width of spatially selective voxels within an area. For this measure of spatial selectivity, we observed an anterior to posterior (rostral-to-caudal) increase of spatial selectivity as well, both in the left hemisphere and right hemisphere (Table 2; Figure 4 B). Specifically, in the left hemisphere, spatial tuning width was broader in the core than in PT, CM, and CL. Finally, spatial tuning width was narrower in PT than in CL (Table 3; Figure 4 B left panel). In the right hemisphere, there was also a difference in spatial tuning width between PT and the core. However, in this hemisphere spatial tuning was sharpest in CM: there was a significant difference between CM and the core , and between CM and CL (Table 3; Figure 4 B right panel).

Next, we investigated cortical inter-area differences in spatial selectivity per behavioral task condition. This revealed that there were differences in the proportion of spatially selective voxels across areas in all behavioral conditions (Table 2). Specifically, post-hoc comparisons revealed that the rostral-to-caudal increase in the proportion of spatially selective voxels was present in all behavioral conditions in the left hemispehere. That is in each condition, there were more spatially selective voxels in PT than in the core and in CM. Further, in the *passive listening* and *sound identification* conditions – but not in the sound localization condition – there were more spatially selective voxels in PT than in CL. In the right hemisphere, we observed significant inter-area differences in the proportion of spatially selective voxels in the *sound identification* condition only. Similar trends were present for the *passive listening* and *sound localization* conditions, but these just failed to reach statistical significance (Table 2). Post-hoc pairwise comparisons for the *sound identification* condition (Table 3) indicate that there are significantly more spatially selective voxels in PT as well as in CM, compared to the core region (see also Figure 5).

We also observed inter-area differences in relative tuning width per behavioral task condition in the left hemisphere. That is, there were significant inter-area differences in all behavioral conditions (Table 2), and in all conditions spatial tuning was sharper in PT than in the core region (see results of post-hoc pairwise comparisons in Table 3). In addition, spatial tuning in PT was sharper than CL in the *passive listening* and *sound identification* condition. Spatial tuning was also sharper in CL than in the core during the *passive listening* and *sound localization* condition. In the right hemisphere, we observed inter-area differences in the *passive listening* and *sound localization* condition (a similar pattern was observed in the *sound identification* condition, but this just failed to reach statistical significance; Table 2). Post-hoc pairwise comparisons show that during *passive listening*, spatial tuning was sharper in PT than in the core region. In addition, spatial tuning was sharper in CM than in either the core region and CL. Also during active *sound localization*, spatial tuning in CM was sharper than in the core and CL, and even PT (Table 3, see also Figure 5).

Task-modulations of spatial selectivity within cortical auditory regions

We then examined, for each cortical area, the effect of task performance on spatial selectivity. There were no differences in the proportion of auditory responsive voxels that were spatially selective across task conditions: none of the cortical regions showed an increase or decrease in the proportion of spatially selective voxels based on task performance (one-tailed Wilcoxon signed-rank tests, all $p > 0.05$, Figure 5 A). However, spatial tuning was sharper in the localization condition compared to the sound identification condition in the left core region (median identification condition = 108.8º, median localization condition = 104.5º, one-tailed Wilcoxon signed-rank test, $p = 0.001$, q[FDR] < 0.05), and in right CM (median identification condition = 91.2º, median localization condition = 85.0º, $p = 0.003$, q[FDR] < 0.05; Figure 5 B). Figure 5C shows the population RAFs, which also reflect the sharpening of spatial selectivity in the left core and right CM during active sound localization.

Next, we investigated the mechanism underlying the observed sharpening of spatial tuning in the left core and right CM during the sound localization condition. Specifically, we evaluated whether the change in spatial tuning between the two active task conditions resulted from response gain (that is, an increase of the BOLD response amplitude at the voxel's preferred location), response sharpening (a decrease of the BOLD response at the voxel's non-preferred location), or a combination of these processes. For this comparison, we defined the voxel's preferred location as the sound azimuth location with the maximum beta value in the task-independent RAF (i.e. the average RAF across the two active task conditions). Similarly, we defined the non-preferred location as the sound azimuth location with the minimum beta value in the average RAF (see also Lee & Middlebrooks, 2011, 2013).

In both cortical areas, the BOLD response at the preferred location was similar for the two active task conditions, while the BOLD response at non-preferred locations was lower in the sound localization than in the sound identification condition. Specifically, Figure 6 shows that the beta values for the preferred location were similar for both active task conditions (reflected by the clustering of beta values around the diagonal; median beta left core in sound identification [sound localization] condition = 0.39 [0.40]; median beta right CM in sound identification [sound localization] condition = 0.27 [0.30]; Wilcoxon signed-rank tests for differences between task conditions, $p > 0.05$). In contrast, the BOLD response at non-preferred locations was lower in the sound localization than in the sound identification condition (most beta values are below the diagonal; median beta left core in sound identification [sound localization] condition = 0.13 [-0.04]; median beta right CM in sound identification [sound localization] condition = 0.04 [-0.11]; Wilcoxon signed-rank tests; left core: $p = 0.002$; right CM: $p = 0.014$; q[FDR] < 0.05). Thus, sharpening of spatial tuning during active sound localization was mainly the result of a decrease of BOLD signal amplitude at non-preferred locations, that is, response sharpening.

Decoding sound azimuth location from fMRI population activity patterns

Next we evaluated whether the encoding of sound azimuth in fMRI activity patterns in the core and in PT varies with behavioral task requirements. Specifically, we applied a population-pattern decoder based on maximum likelihood estimation (MLE) to the measured fMRI responses to the probe sounds in the sound identification and sound localization condition (see *Methods*). Figure 7 shows for each cortical area and task condition the absolute error of the population pattern decoder as a function of sound azimuth location. There was no difference in decoding performance between ipsi- and contralateral locations: a comparison of the average absolute error between hemifields (i.e. the average absolute error across -30º, -60º, and -90º, versus the average across +30º, +60º, and +90º) did not yield significant results either for the core or for PT, in any behavioral task condition (two-sided Wilcoxon signed-rank test per area and task condition, FDR corrected for multiple comparisons, all q[FDR] > 0.05).

For the purpose of statistical comparisons between cortical areas and behavioral task conditions, we computed the average absolute error across azimuth positions for each area and task condition. Figure 7 B shows that the population pattern decoder performed better than chance level in the left and right core in the sound localization condition. That is, in these areas and task conditions the absolute error was significantly lower than chance (one-sided Wilcoxon signed-rank test, FDR corrected for multiple corrections; median absolute error sound localization condition left core = 61.1º, right core = 62.1º, chance error = 68.6º, p = 0.009 for both regions, q[FDR] < 0.05). Chance level was computed with a permutation testing procedure in which we randomly scrambled the RAFs of each participant (1500 iterations). In left PT, the pattern decoder also performed better than chance in the localization condition (median absolute error left PT = 58.9º, p = 9.8E-4, q[FDR] < 0.05). Similarly, in right PT the pattern decoder performed marginally better than chance in the localization condition (median absolute error right PT = 60.0º, p = 0.051, q[FDR] = 0.076). However, in the sound identification condition the absolute error was larger than chance level in all cortical areas (median absolute error for

519 the sound identification condition per area: left core = 75.0º, right core = 66.4º, left PT = 70.7º,
520 right PT = 71.8º, $p > 0.05$; Figure 7 B), indicating that the pattern decoder did not perform well
521 for this behavioral condition.

522 We then tested for differences in sound location decoding performance for the probe sounds
523 between task conditions, within each cortical area. This showed that the pattern decoder
524 performed significantly better in the sound localization than in the sound identification condition
525 in the left core region – that is, the absolute error was significantly lower (one-sided Wilcoxon
526 signed-rank test, FDR-corrected for multiple comparisons; $p = 0.003$, q[FDR] < 0.05; Figure 7
527 B). In left PT we observed a similar task effect, but this did not reach statistical significance ($p =$
528 0.04, q[FDR] = 0.1). Figure 7 A shows that the absolute error decreased especially at the
529 midline and in contralateral space (0º to + 90º) for both the core and PT in the left hemisphere.
530 There was no significant effect of task in the right core or in right PT ($p > 0.05$; Figure 7). For the
531 right core, this may be a consequence of the relatively high performance of the pattern decoder
532 in the sound identification condition. In particular, sound azimuth location estimates were
533 significantly more accurate in the right, than in the left core in the sound identification condition
534 (two-sided Wilcoxon signed-rank test; $p = 0.022$, q[FDR] < .05), but not in the sound localization
535 condition ($p > 0.05$; Figure 7 B),

536 We also tested for each task condition whether there was a difference in decoding accuracy
537 between cortical areas. In the left hemisphere, the absolute error was lower in PT than in the
538 core region in the sound identification condition ($p = 0.0098$, q[FDR] < 0.05) but not in the sound
539 localization condition ($p > 0.05$). Figure 7 A shows that the inter-area difference in the sound
540 identification condition was mainly a result of lower absolute errors in PT in peripheral space. In
541 the right hemisphere, there was no significant difference between the core and PT either in the
542 sound identification condition ($p > 0.05$) or in the sound localization condition ($p > 0.05$). Note
543 that the lower absolute error observed in left PT was not a consequence of a larger number of

voxels in this cortical region: the inter-area effect persisted even if the number of voxels in PT included in the analysis was matched to the number of voxels in the core region (see *Methods* and Figure 7 C).

Finally, we applied the maximum-likelihood decoder to the fMRI activity patterns of the left and right hemisphere together: we provided the data of both hemispheres combined as input for the pattern decoder. Note that to ensure that the number of voxels on which the pattern decoder operates does not influence the sound location estimates, we randomly sampled half of the voxels in the relevant region within a hemisphere and combined this with a random sample of half of the voxels in the other hemisphere. This procedure was repeated 200 times, and we computed the absolute error of the two-hemisphere decoder as the average absolute error across those 200 iterations.

Figure 8 shows that combining the activity patterns in the two hemispheres resulted in lower absolute errors when decoding azimuth position for probe sounds in the sound identification, but not for probe sounds in the sound localization condition. Specifically, absolute error scores were lower than chance level in the sound identification condition in both the core and in PT (median absolute error core = 62.4º, median absolute error PT = 59.3º, chance error = 68.8º, $p = 0.03$ and $p = 0.009$ respectively, q[FDR] < 0.05). In addition, the absolute error in PT was lower for the combined data than for either the left PT only ($p = 0.016$, q[FDR] < 0.05), or the right PT only ($p = 0.003$, q[FDR] < 0.05). Inspecting absolute error as a function of sound azimuth location (Figure 8 A), shows that combining the data of left and right PT resulted in lower absolute error scores mainly in the periphery (-90º, -60º, +60º, and +90º). In contrast, for the core the combination of the data of the left and right hemisphere resulted in more accurate azimuth estimates in comparison to the left core ($p = 0.002$), but not in comparison to the right core ($p > 0.05$). Further, the absolute error as a function of sound azimuth position (Figure 8 A) shows that the absolute errors resulting from the combined data were similar to those resulting

from the decoder operating on the right core only. This indicates that the azimuth estimates resulting from the pattern decoder operating on the core in two hemispheres are driven by the activity patterns in the right core, rather than showing an improvement larger than the available information in either core.

DISCUSSION

The major findings of the present study are that spatial selectivity of the left primary auditory core cortex and right area CM are dynamic and dependent on behavioral requirements, that fMRI activity patterns in the left core carry more information on sound azimuth location when participants engage in a sound-localization task (in comparison to a task unrelated to sound localization), and that integrating fMRI activity patterns measured during a 'what' task – but not during a 'where' task – across bilateral PT results in more accurate sound azimuth location estimates than in either left or right PT separately. Together, these results highlight the adaptive potential of spatial tuning in the primary auditory cortex based on behavioral demands. A possible mechanism for the observed task-modulation of spatial sensitivity in primary auditory cortex is the feedback from functionally specialized regions (planum temporale) to this cortical area. Specifically, such feedback connections from higher-order to primary regions may be modulated by behavioral requirements to enable dynamic spatial sensitivity in the latter. Finally, these findings provide new insights into models of sound location encoding in unilateral and bilateral human auditory cortex.

Dynamic spatial tuning in human auditory cortex

Posterior auditory cortical regions are thought to be part of a functionally specialized stream for sound location processing in animals (Harrington et al., 2008; Lomber & Malhotra, 2008; Stecker et al., 2005; Stecker & Middlebrooks, 2003; Tian et al., 2001) and humans (Ahveninen et al., 2006; Alain, Arnott, Hevenor, Graham, & Grady, 2001; Arnott et al., 2004; Brunetti et al.,

2005; Deouell et al., 2007; Derey et al., 2015). While we replicate these inter-area differences in spatial selectivity between primary core and higher-order areas – and specifically the advantage of caudal belt regions – that have been reported previously for passive listening or non-spatial task-conditions, we also show that these differences are reduced in the left core and right CM when humans engage in an active sound localization task. Thus, our findings indicate that, depending on the behavioral requirements, primary auditory areas may contribute to sound location processing as well.

Such task-dependent modulations of spatial sensitivity have not previously been observed in humans. Zimmer and Macaluso (2005) reported a relationship between the level of activity in posterior auditory regions and successful sound localization, but did not investigate cortical spatial selectivity. Further, a recent neuroimaging study in humans did not report a modulation of either ILD or ITD selectivity based on task-performance (Higgins et al., 2017). Yet, in the latter study, the authors considered binaural cue response functions averaged across all auditory responsive voxels within the auditory cortex, which may have diluted the results. That is, our analyses show that task modulations of spatial selectivity are localized specifically in the left core and right CM.

Our findings in human auditory cortex are compatible with animal studies showing that the performance of both spatial and non-spatial tasks affects neuronal receptive fields in primary auditory cortex (Fritz, Shamma, Elhilali, & Klein, 2003; Lee & Middlebrooks, 2011; Otazu, Tai, Yang, & Zador, 2009). One hypothesis is that higher-order, functionally specialized cortical areas such as PT modulate spatial tuning in primary auditory cortex via back-projections. In particular, our data are compatible with theoretical frameworks of sensory processing such as the “reverse hierarchy” (Ahissar et al., 2009) and recurrent processing models (Bullier, 2001; Lamme & Roelfsema, 2000). Similar to visual cortex, the auditory cortex is characterized by dense reciprocal connections between primary and higher-order cortical areas (Kaas & Hackett,

2000; Lee & Winer, 2011). Lateral prefrontal cortex (PFC) may mediate such feedback processing: lateral PFC is known to project back to early regions of the lateral auditory belt (Romanski et al., 1999) and has been implicated in a two-stage model of categorization of sounds (Jiang, Chevillet, Rauschecker, & Riesenhuber, 2018).

Differences in sound location processing between the left and right auditory pathway

In humans, lesion and functional imaging studies suggest that the right (sub)cortical pathway may contain a representation of the entire acoustic azimuth, while in the left (sub)cortical pathway the representation of the contralateral acoustic azimuth is thought to be pre-dominant (e.g. Briley, Kitterick, & Summerfield, 2013; Higgins et al., 2017; Krumbholz et al., 2005; Spierer, Bellmann-Thiran, Maeder, Murray, & Clarke, 2009; Zatorre & Penhune, 2001). Differential spatial processing between the left and right auditory pathway has also been observed in several animal species. For instance, Day and Delgutte (2013) observed in rabbit inferior colliculus a gradient of deteriorating sound location decoding accuracy from locations at the midline towards the periphery. In contrast, in monkeys, Miller and Recanzone (2009) observed in area A1 and CL most accurate sound location decoding results in contralateral space, with low decoding accuracies at the midline and especially in ipsilateral space: the magnitude of sound location estimation errors in the ipsilateral hemifield and around the midline was distinctly higher than the errors observed in the present study. Only in area R were decoding errors lower around the midline than in either ipsi- or contralateral space. Here we did not observe a difference in location decoding accuracy between ipsi- and contralateral space either for the left or right auditory cortex. Yet, our results did reflect sharper spatial tuning in the right than left core when the task was unrelated to sound location (the ‘what’ task), which may be a reflection of the hypothesized right dominance for human spatial hearing. Future research with non-invasive lesion techniques in humans combined with advanced neuroimaging and

642 computational modeling studies is required to elucidate these potential differences between the
643 left and right human auditory pathway.

644 **Integrating information on sound azimuth location across hemispheres**

645 Our results show that the integration of sound location processing across the two hemispheres
646 may be task-dependent. Specifically, location estimates based on fMRI activity patterns in
647 bilateral PT were more accurate than those based on either left or right PT independently for the
648 task condition unrelated to sound localization ('what' task), while, this bilateral advantage was
649 not present during active localization ('where' task). For the core region, we also observed a
650 bilateral advantage for the 'what' task compared to the left core separately, but not for the right
651 core. This suggests that the bilateral advantage, is merely a reflection of the more accurate
652 decoding obtained for the right core in itself. Similar to PT, no bilateral decoding improvement
653 was observed during active sound localization for the core region. Thus, fMRI activity patterns in
654 left and right PT - and possibly in the left and right core - contain complementary information on
655 sound azimuth location when participants are not engaged in active sound localization, resulting
656 in better location estimates when the information in the two hemispheres is combined. In
657 contrast, information in the two hemispheres appears to be overlapping during active sound
658 localization, such that combining the information across the hemispheres appears to be
659 redundant during this behavioral condition.

660 This may be explained by a task-dependent strength of functional callosal connections. In
661 particular, in macaques there are major interhemispheric connections both between the left and
662 right core, and between left and right parabelt (Kaas & Hackett, 2000). If similar callosal
663 connections between bilateral primary and higher order auditory cortices exist in humans, it is
664 conceivable that during active sound localization the functional connectivity between left and
665 right PT increases compared to during non-localization tasks. As a consequence, spatial

666 processing in left PT may modulate spatial processing in right PT during active localization (and
667 vice versa), while spatial information in left and right PT is relatively independent – and thus
668 complementary – during non-spatial tasks. Alternatively, corticofugal projections (e.g. Winer &
669 Schreiner, 2005) may strengthen during active sound localization, and thereby indirectly
670 modulate sound location processing in the contralateral hemisphere.

671 The observed task-dependency of bilateral integration of information is also of interest for the
672 ongoing debate about the computational mechanisms underlying sound location processing in
673 mammals. In particular, models for neural population coding of sound azimuth location have
674 received wide attention in recent years, including population coding within a single hemisphere
675 (unilateral population coding, e.g. Day & Delgutte, 2013; Miller & Recanzone, 2009), *unilateral*
676 *opponent population coding* based on two oppositely tuned channels within a single hemisphere
677 (i.e. an ipsi- and a contralaterally tuned channel, Stecker et al., 2005), and *bilateral opponent*
678 population coding based on combining the sound azimuth information of contralaterally tuned
679 channels in each hemisphere (e.g. Derey et al., 2015; McAlpine et al., 2001; Ortiz-Rios et al.,
680 2017). Our current results suggest that the degree to which information is combined across
681 hemispheres may be dependent on behavioral requirements, indicating that unilateral and
682 bilateral models of sound location encoding may not be mutually exclusive.

683

References

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1515), 285-299.

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., . . . Shinn-Cunningham, B. G. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proceedings of the National Academy of Sciences, 103*(39), 14608-14613.

Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences, 98*(21), 12301-12306.

Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage, 22*(1), 401-408.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300.

Briley, P. M., Kitterick, P. T., & Summerfield, A. Q. (2013). Evidence for opponent process analysis of sound source location in humans. *Journal of the Association for Research in Otolaryngology, 14*(1), 83-101.

Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Della Penna, S., Ferretti, A., . . . Tartaro, A. (2005). Human brain activation during passive listening to sounds from different locations: an fMRI and MEG study. *Human brain mapping, 26*(4), 251-261.

Bullier, J. (2001). Integrated model of visual processing. *Brain research reviews, 36*(2-3), 96-107.

Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S., Clarke, S., & Saenz, M. (2011). Human primary auditory cortex follows the shape of Heschl's gyrus. *Journal of Neuroscience, 31*(40), 14067-14075.

Day, M. L., & Delgutte, B. (2013). Decoding sound source location and separation using neural population activity patterns. *The Journal of Neuroscience, 33*(40), 15837-15847.

De Martino, F., Moerel, M., van de Moortele, P.-F., Ugurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2013). Spatial organization of frequency preference and selectivity in the human inferior colliculus. *Nature communications, 4*, 1386.

Deouell, L. Y., Heller, A. S., Malach, R., D'Esposito, M., & Knight, R. T. (2007). Cerebral responses to change in spatial location of unattended sounds. *Neuron, 55*(6), 985-996.

Derey, K., Valente, G., de Gelder, B., & Formisano, E. (2015). Opponent Coding of Sound Location (Azimuth) in Planum Temporale is Robust to Sound-Level Variations. *Cerebral Cortex*, bhv269.

Formisano, E., Kim, D.-S., Di Salle, F., van de Moortele, P.-F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron, 40*(4), 859-869.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature neuroscience, 6*(11), 1216.

Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human brain mapping, 27*(5), 392-401.

Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature, 430*(7000), 682.

Harrington, I. A., Stecker, G. C., Macpherson, E. A., & Middlebrooks, J. C. (2008). Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hearing research, 240*(1-2), 22-41.

729 Higgins, N. C., McLaughlin, S. A., Rinne, T., & Stecker, G. C. (2017). Evidence for cue-independent spatial
730 representation in the human auditory cortex during active listening. *Proceedings of the National*
731 *Academy of Sciences*, 201707522.
732 Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural
733 populations. *Nature neuroscience, 9*(5), 690-696.
734 Jiang, X., Chevillet, M.A., Rauschecker, J.P., & Riesenhuber, M. (2018). Training humans to categorize
735 monkey calls: auditory feature and category selective neural tuning changes. *Neuron*, *98*(2), 405-
736 416.
737 Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates.
738 *Proceedings of the National Academy of Sciences, 97*(22), 11793-11799.
739 Kim, J.J., Crespo-Facorro, B., Andreasen, N. C., O'Leary, D. S., Zhang, B., Harris, G., & Magnotta, V. A.
740 (2000). An MRI-based parcellation method for the temporal lobe. *Neuroimage, 11*(4), 271-288.
741 King, A. J., Bajo, V. M., Bizley, J. K., Campbell, R. A., Nodal, F. R., Schulz, A. L., & Schnupp, J. W. (2007).
742 Physiological and behavioral studies of spatial coding in the auditory cortex. *Hearing research,*
743 *229*(1), 106-115.
744 Krumbholz, K., Schönwiesner, M., von Cramon, D. Y., Rübsamen, R., Shah, N. J., Zilles, K., & Fink, G. R.
745 (2005). Representation of interaural temporal information from left and right auditory space in
746 the human planum temporale and inferior parietal lobe. *Cerebral Cortex, 15*(3), 317-324.
747 Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and
748 recurrent processing. *Trends in neurosciences, 23*(11), 571-579.
749 Leaver, A. M., & Rauschecker, J. P. (2016). Functional Topography of Human Auditory Cortex. *The*
750 *Journal of Neuroscience, 36*(4), 1416-1428.
751 Lee, C. C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task
752 performance. *Nature neuroscience, 14*(1), 108-114.
753 Lee, C. C., & Middlebrooks, J. C. (2013). Specialization for sound localization in fields A1, DZ, and PAF of
754 cat auditory cortex. *Journal of the Association for Research in Otolaryngology, 14*(1), 61-82.
755 Lee, C. C., & Winer, J. A. (2011). A synthesis of auditory cortical connections: thalamocortical,
756 commissural and corticocortical systems. In *The auditory cortex* (pp. 147-170): Springer.
757 Lomber, S. G., & Malhotra, S. (2008). Double dissociation of'what'and'where'processing in auditory
758 cortex. *Nature neuroscience, 11*(5), 609-616.
759 McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in
760 mammals. *Nature neuroscience, 4*(4), 396-401.
761 McLaughlin, S. A., Higgins, N. C., & Stecker, G. C. (2016). Tuning to binaural cues in human auditory
762 cortex. *Journal of the Association for Research in Otolaryngology, 17*(1), 37-53.
763 Miller, L. M., & Recanzone, G. H. (2009). Populations of auditory cortical neurons can accurately encode
764 acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences,*
765 *106*(14), 5931-5935.
766 Moerel, M., De Martino, F., & Formisano, E. (2012). Processing of natural sounds in human auditory
767 cortex: tonotopy, spectral tuning, and relation to voice sensitivity. *The Journal of Neuroscience,*
768 *32*(41), 14205-14216.
769 Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human
770 auditory cortical areas. *Frontiers in neuroscience, 8*, 225.
771 Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of*
772 *Experimental Psychology: Human Perception and Performance, 21*(2), 387.
773 Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human
774 primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference
775 system. *Neuroimage, 13*(4), 684-701.

776 Ortiz-Rios, M., Azevedo, F. A., Kuśmierek, P., Balla, D. Z., Munk, M. H., Keliris, G. A., . . . Rauschecker, J. P.
777 (2017). Widespread and Opponent fMRI Signals Represent Sound Location in Macaque Auditory
778 Cortex. *Neuron, 93*(4), 971-983. e974.
779 Otazu, G. H., Tai, L.-H., Yang, Y., & Zador, A. M. (2009). Engaging in an auditory task suppresses
780 responses in auditory cortex. *Nature neuroscience, 12*(5), 646.
781 Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates
782 illuminate human speech processing. *Nature neuroscience, 12*(6), 718-724.
783 Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in
784 auditory cortex. *Proceedings of the National Academy of Sciences, 97*(22), 11800-11806.
785 Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque
786 nonprimary auditory cortex. *SCIENCE-NEW YORK THEN WASHINGTON-*, 111-111.
787 Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999). Dual
788 streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature*
789 *neuroscience, 2*(12), 1131.
790 Rorden, C., & Driver, J. (2001). Spatial deployment of attention within and across hemifields in an
791 auditory task. *Experimental Brain Research, 137*(3-4), 487-496.
792 Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous
793 mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3),
794 555.
795 Spierer, L., Bellmann-Thiran, A., Maeder, P., Murray, M. M., & Clarke, S. (2009). Hemispheric
796 competence for auditory spatial representation. *Brain, 132*(7), 1953-1966.
797 Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural
798 populations in the auditory cortex. *PLoS Biol, 3*(3), e78.
799 Stecker, G. C., & Middlebrooks, J. C. (2003). Distributed coding of sound locations in the auditory cortex.
800 *Biological cybernetics, 89*(5), 341-349.
801 Striem-Amit, E., Hertz, U., & Amedi, A. (2011). Extensive cochleotopic mapping of human auditory
802 cortical fields obtained with phase-encoding FMRI. *PLoS One, 6*(3), e17832.
803 Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic
804 organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal*
805 *of neurophysiology, 91*(3), 1282-1296.
806 Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus
807 monkey auditory cortex. *Science, 292*(5515), 290-293.
808 Tournoux, T., & Talairach, J. (1988). Co-planar stereotaxic atlas of the human brain. *Theime, Stuggart,*
809 *Germany*.
810 Van der Zwaag, W., Gentile, G., Gruetter, R., Spierer, L., & Clarke, S. (2011). Where sound position
811 influences sound object representations: a 7-T fMRI study. *Neuroimage, 54*(3), 1803-1811.
812 Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch
813 sequences in the human auditory brain. *Journal of Neuroscience, 23*(13), 5799-5804.
814 Wessinger, C., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., & Rauschecker, J. (2001). Hierarchical
815 organization of the human auditory cortex revealed by functional magnetic resonance imaging.
816 *Journal of cognitive neuroscience, 13*(1), 1-7.
817 Winer, J. A., & Schreiner, C. E. (2005). The central auditory system: a functional analysis. In *The inferior*
818 *colliculus* (pp. 1-68): Springer.
819 Zatorre, R. J., & Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex.
820 *Journal of Neuroscience, 21*(16), 6321-6328.
821 Zimmer, U., & Macaluso, E. (2005). High binaural coherence determines successful sound localization
822 and increased activity in posterior auditory areas. *Neuron, 47*(6), 893-905.

823

FIGURE LEGENDS

Figure 1. Stimuli. (A) Azimuth locations at which sound sources were presented. (B) Example of a probe trial (top), a target trial for the sound localization task (middle), and a target trial for the sound identification task (bottom). A probe trial consisted of a block of five stimulus presentations at one azimuth location. In the sound localization task, the target trial consisted of five stimulus presentations as well, yet for the fourth (depicted here) or fifth repetition the azimuth location was changed. For target trials in the sound identification condition, azimuth location remained constant across the five stimulus repetitions but the fourth or fifth repetition was replaced by a deviant click train. (C) Lines reflect the interaural time difference (ITD; left) and interaural level difference (ILD; right) for stimuli at a specific sound azimuth position, averaged across the binaural recordings of all participants. ILD was computed as the arithmetic difference in power (measured as root mean square [RMS]) between the left and right channel of each binaural recording. To compute ITD, we first computed the interaural phase difference (IPD) which we subsequently converted to time differences. (D) Plotted is the power spectrum of the left channel of the binaural recordings (i.e. the left ear) at specific azimuth positions, averaged across all participants. The difference in power in specific frequency bands dependent on sound azimuth location illustrates the availability of spectral cues in the recordings. Colors similar to (C).

Figure 2. Estimating sound azimuth location with a maximum-likelihood population pattern decoder. Small graphs show the log-likelihood function for each voxel for a given sound azimuth location (rows), with the fMRI response strength (beta value) on the-x axis, and the log-likelihood on the y-axis. Large graph on the right shows the resulting population log-likelihood function, which is the sum of the log-likelihood functions of the individual voxels at each location.

Figure 3. Parcellation of the human auditory cortex. (A) The figure shows an enlarged view of the superior temporal plane in the right hemisphere, with a schematic overview of the parcellation used in the present study overlaid on top. (B) Figures show the left and right superior temporal plane of a representative participant with the group map of frequency preference overlaid (top row; warm colors indicate a maximum response to low frequencies, cold colors to high frequencies), and frequency selectivity (bottom row; orange to green colors indicate broad tuning, blue to purple colors indicate progressively sharper tuning. (C) Similar to (A) but for a single representative participant.

Figure 4. Spatial selectivity across auditory cortical areas in humans. (A) Box-plots show, for each cortical area, the distribution of the proportion of spatially selective voxels across participants (averaged across task conditions). (B) Box-plots reflect the distribution of relative spatial tuning width (ERRF width, averaged across task conditions) across participants. The central circle of a box indicates the median of the distribution, the edges the 25^{th} and 75^{th} percentiles, and lines the full range of values. Circles represent outliers. Horizontal lines indicate a significant difference between areas at $p < 0.05$, FDR corrected for multiple comparisons at $q < 0.05$. CM = caudo-medial region. CL = caudo-lateral region. PT = planum temporale.

Figure 5. Task modulations of spatial selectivity in human auditory cortex. (A) Box-plots show for each task condition the distribution of the proportion of voxels that exhibit a spatially selective response across participants. Black boxes indicate the passive listening condition, red boxes the sound identification condition, and blue boxes the sound localization condition. (B) Box-plots reflect the distribution of relative spatial tuning width (ERRF width) across participants for each area and task condition. Colors similar to (A). The central circle of a box indicates the median of the distribution, the edges the 25^{th} and 75^{th} percentiles, and lines the full range of values. Circles represent outliers. Horizontal lines indicate a significant difference between areas at $p < 0.05$, FDR corrected for multiple comparisons at $q < 0.05$. (C) Population RAFs are

plotted for the spatially selective voxels within an area for the two active task conditions. RAFs are averaged across participants, blue lines indicate the sound identification condition, red lines the sound localization condition. CM = caudo-medial region. CL = caudo-lateral region. PT = planum temporale.

Figure 6. Sharper spatial selectivity during active sound localization is a result of response sharpening. Scatterplots show for each participant the average beta value across voxels that exhibited sharper spatial selectivity (i.e. a decrease in ERRF width of 15% or more) during the sound localization condition (y-axis) than sound identification condition (x-axis). at the preferred (filled circles) and non-preferred location (open circles) for the left core region (left panel) and right CM (right panel). Circles below the diagonal reference line reflect a decrease in beta value in the sound localization condition.

Figure 7. Decoding sound azimuth from population pattern activity in the core region and PT during a sound identification ('what') and a sound localization ('where') task. (A) Lines reflect the average absolute error of the sound azimuth estimate resulting from the population pattern decoder (y axis) as a function of actual sound azimuth (x axis) for a particular cortical area and task condition. Light blue lines: core region during sound identification task. Dark blue lines: core region during sound localization task. Light green lines: PT during sound identification task. Dark green lines: PT during sound localization task. Error bars reflect the standard error of the mean (SEM). (B) Box-plots of the absolute error of the sound azimuth estimates averaged across the seven sound azimuth positions. Colors similar to (A). Horizontal black lines at the top of the figure indicate a significant difference in prediction error between cortical areas or task conditions ($p < 0.05$, q[FDR] < 0.05). Horizontal red lines at the bottom of the figure indicate that the absolute error is below chance level ($p < 0.05$, q[FDR] < 0.05). (C) Lines reflect the performance of the population pattern decoder for PT controlled for the number of voxels. Similar to (A), lines reflect the average absolute error. Solid lines are identical to those

for area PT depicted in (A). Dashed lines show the average absolute error across random samples (200 iterations) of voxels in PT. Specifically, for each participant we sampled a number of voxels from PT equal to the number of voxels included in the analysis for the core. Error bars reflect the standard error of the mean (SEM).

Figure 8. Decoding sound azimuth from population pattern activity across two hemispheres. (A) Lines reflect the average absolute error of the sound azimuth estimate resulting from the population pattern decoder (y axis) as a function of actual sound azimuth (x axis) for a particular cortical area and task condition. Light blue lines: core region during sound identification task. Dark blue lines: core region during sound localization task. Light green lines: PT during sound identification task. Dark green lines: PT during sound localization task. Error bars reflect the standard error of the mean (SEM). (B) Box-plots of the absolute error of the sound azimuth estimates averaged across the seven sound azimuth positions. Colors similar to (A). Gray boxes are identical to the boxes shown in Figure 6 and show the absolute error for the left hemisphere only (left-most gray box) and for the right hemisphere only (right-most gray box) for comparison. Horizontal black lines at the top of the figure indicate a significant difference in prediction error between cortical areas or task conditions ($p < 0.05$, q[FDR] < 0.05. Horizontal red lines at the bottom of the figure indicate that the absolute error is below chance level ($p < 0.05$, q[FDR] < 0.05).

TABLES

Table 1. Number of auditory responsive voxels per cortical area.

	Average number of voxels (standard deviation)	
Cortical area	**Left hemisphere**	**Right hemisphere**
Core	160.0 (24.2)	138.9 (19.1)
CM	86.6 (14.9)	83.6 (12.9)
CL	73.9 (15.5)	68.3 (12.2)
RM	56.0 (15.0)	51.6 (15.7)
RL	57.8 (12.3)	44.6 (16.7)
PT	271.3 (49.4)	238.4 (48.0)

Table shows number of auditory responsive voxels for each cortical area, averaged across participants. CM = caudo-medial area. CL = caudo-lateral area. RM = rostro-medial area. RL = rostro-lateral area. PT = planum temporale.

Table 2. Differences in the proportion of spatially selective voxels and tuning width between cortical auditory areas. Table shows the statistical results of Friedman tests to compare the proportion of spatially selective voxels (upper part) and relative tuning width (lower part) between cortical auditory areas, as well as the median value for each cortical area.

		Left hemisphere							Right hemisphere						
		Friedman test			*Median*				*Friedman test*			*Median*			
	Condition	**X^2**	***df***	***p***	**Core**	**CM**	**CL**	**PT**	**X^2**	***df***	***p***	**Core**	**CM**	**CL**	**PT**
Proportion spatially selective voxels	Passive	18.6	3	0.0003*	51.2	63.6	65.6	68.8	7.1	3	0.067	58.3	62.2	55.9	66.1
	Identification	18.6	3	0.0003*	50.0	58.8	61.2	66.4	10.1	3	0.018*	51.3	60.7	60.8	63.3
	Localization	9.4	3	0.024*	52.3	55.3	54.1	63.9	7.9	3	0.048	56.2	59.3	58.2	64.7
	Average	19.5	3	0.0002*	50.6	56.3	60.6	66.5	12.4	3	0.0062*	53.8	62.6	56.6	65.2
Tuning width	Passive	15.9	3	0.0012*	108.3	97.7	96.8	90.7	17.9	3	0.0005*	102.0	90.5	100.6	94.0
	Identification	15.6	3	0.0014*	108.8	95.6	95.5	91.8	7.4	3	0.058	103.1	91.2	97.4	97.3
	Localization	10.1	3	0.018*	104.5	98.8	99.2	92.8	15.4	3	0.0015*	102.0	85.0	100.4	95.1
	Average	16.7	3	0.0008*	105.3	95.9	98.7	91.2	18.4	3	0.0004*	101.6	89.1	98.6	95.7

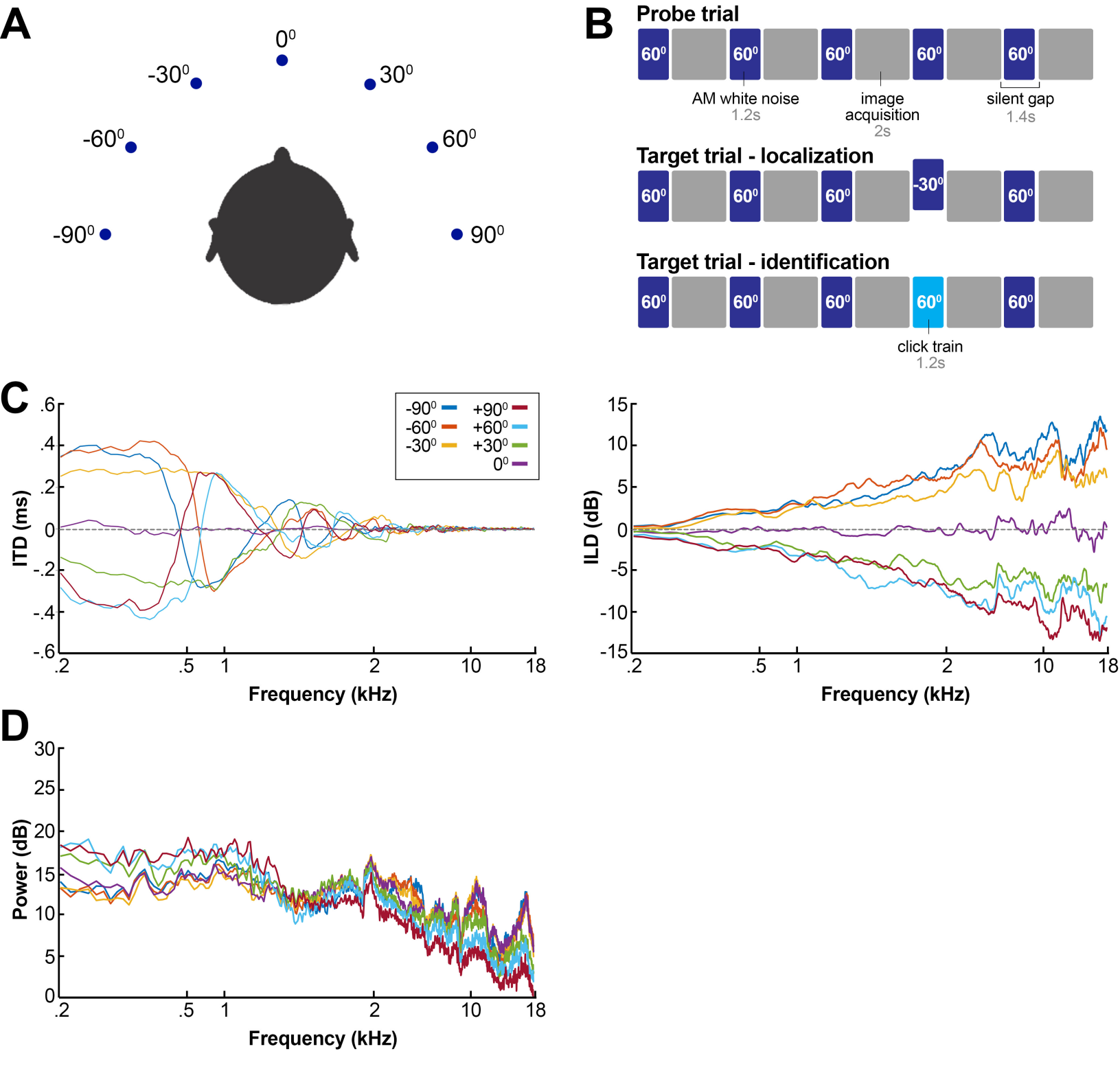
* *p* value significant after FDR correction for multiple comparisons ($q < 0.05$). Proportions are displayed in percentages, tuning width in degrees. CM = caudomedial area; CL = caudolateral area; PT = planum temporale.

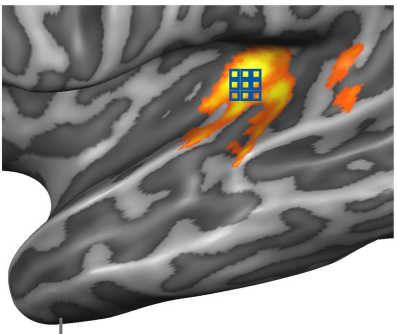
930 **Table 3.** Statistical results (*p* values) of post-hoc pairwise comparisons of the proportion of
931 spatially selective voxels (upper part) and tuning width (lower part) between cortical auditory
932 regions (two-sided Wilcoxon signed-rank tests).

			Left hemisphere			Right hemisphere		
			Core	CM	CL	Core	CM	CL
Proportion of spatially selective voxels	Passive	CM	0.320			n.a.		
		CL	0.001*	0.206		n.a.	n.a.	
		PT	0.001*	0.024*	0.320	n.a.	n.a.	n.a.
	Identification	CM	0.206			0.007*		
		CL	0.005*	0.175		0.577	0.042	
		PT	0.001*	0.002*	0.054	0.010*	0.966	0.054
	Localization	CM	0.638			n.a.		
		CL	0.577	0.700		n.a.	n.a.	
		PT	0.002*	0.010*	0.032	n.a.	n.a.	n.a.
	Average	CM	0.320			0.042		
		CL	0.002*	0.175		0.700	0.014*	
		PT	0.001*	0.005*	0.014*	0.007*	0.465	0.010*
Tuning width	Passive	CM	0.042			0.001*		
		CL	0.005*	0.765		0.320	0.003*	
		PT	0.001*	0.083	0.024*	0.019*	0.148	0.067
	Identification	CM	0.007*			n.a.		
		CL	0.010*	0.638		n.a.	n.a.	
		PT	0.001*	0.278	0.032*	n.a.	n.a.	n.a.
	Localization	CM	0.067			0.005*		
		CL	0.083	0.638		0.465	0.005*	
		PT	0.002*	0.465	0.067	0.042	0.019*	0.413
	Average	CM	0.019*			0.002*		
		CL	0.005*	0.700		0.148	0.001*	
		PT	0.001*	0.123	0.032*	0.019*	0.054	0.365

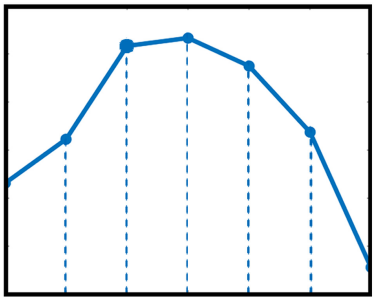
933 * Significant after FDR correction at $q < 0.05$. Note that we only performed post-hoc pairwise
934 comparisons for those behavioral conditions that showed significant inter-area differences as
935 assessed with the Friedman tests displayed in Table 2. CM = caudomedial area; CL =
936 caudolateral area; PT = planum temporale.

937


A
0⁰
-30⁰
30⁰
-60⁰
60⁰
-90⁰
90⁰
B
Probe trial
60⁰
AM white noise
1.2s
image
acquisition
2s
silent gap
1.4s
Target trial - localization
-30⁰
Target trial - identification
click train
1.2s
C
-90⁰ +90⁰
-60⁰ +60⁰
-30⁰ +30⁰
0⁰
ITD (ms)
ILD (dB)
Frequency (kHz)
D
Power (dB)
Frequency (kHz)



Population log likelihood

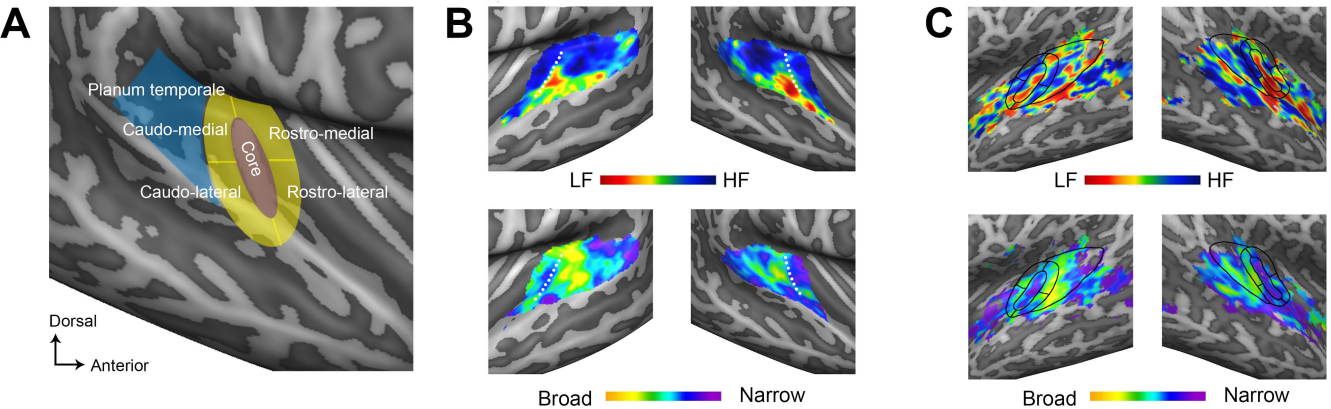


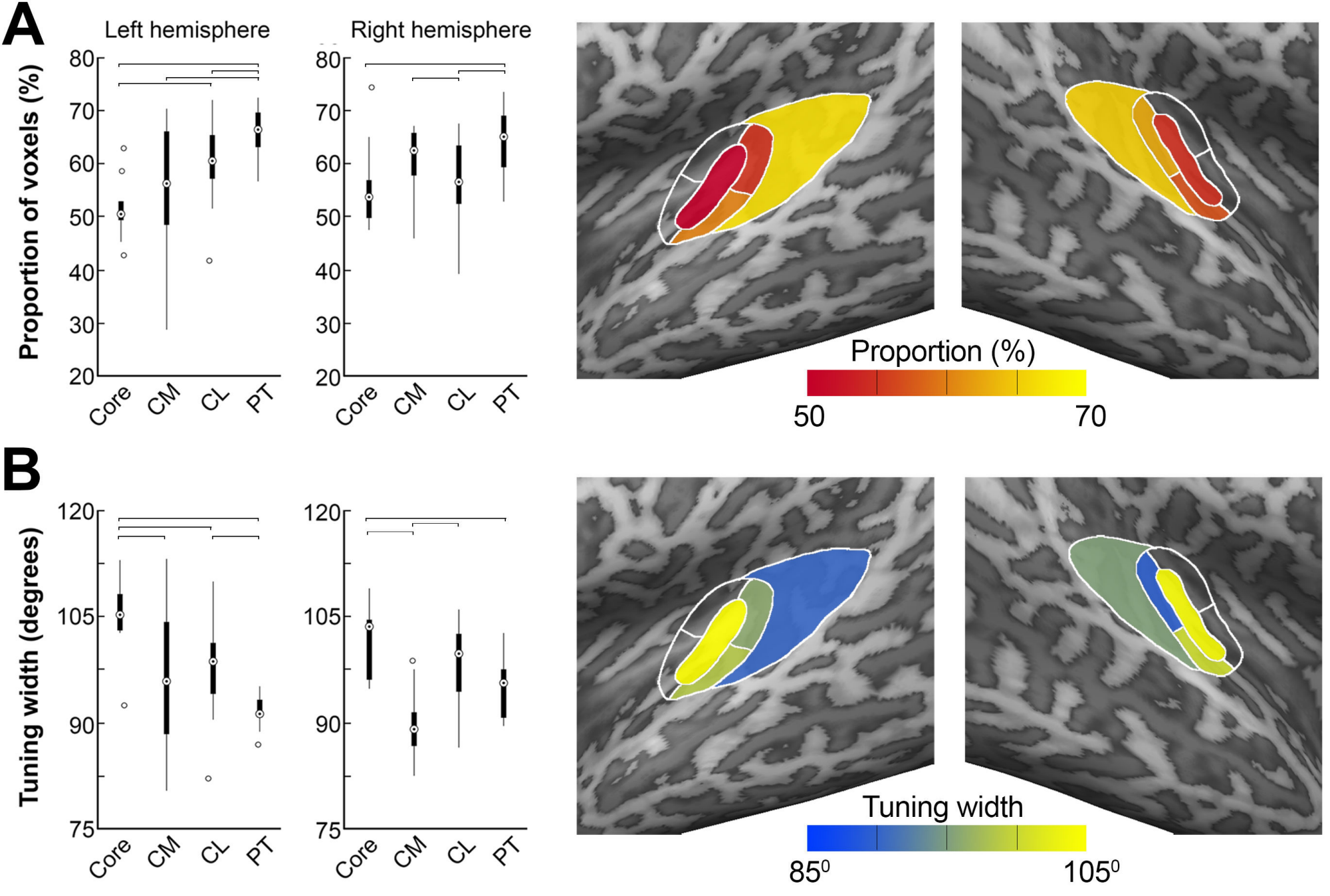
Log likelihoods for azimuth **-90⁰**

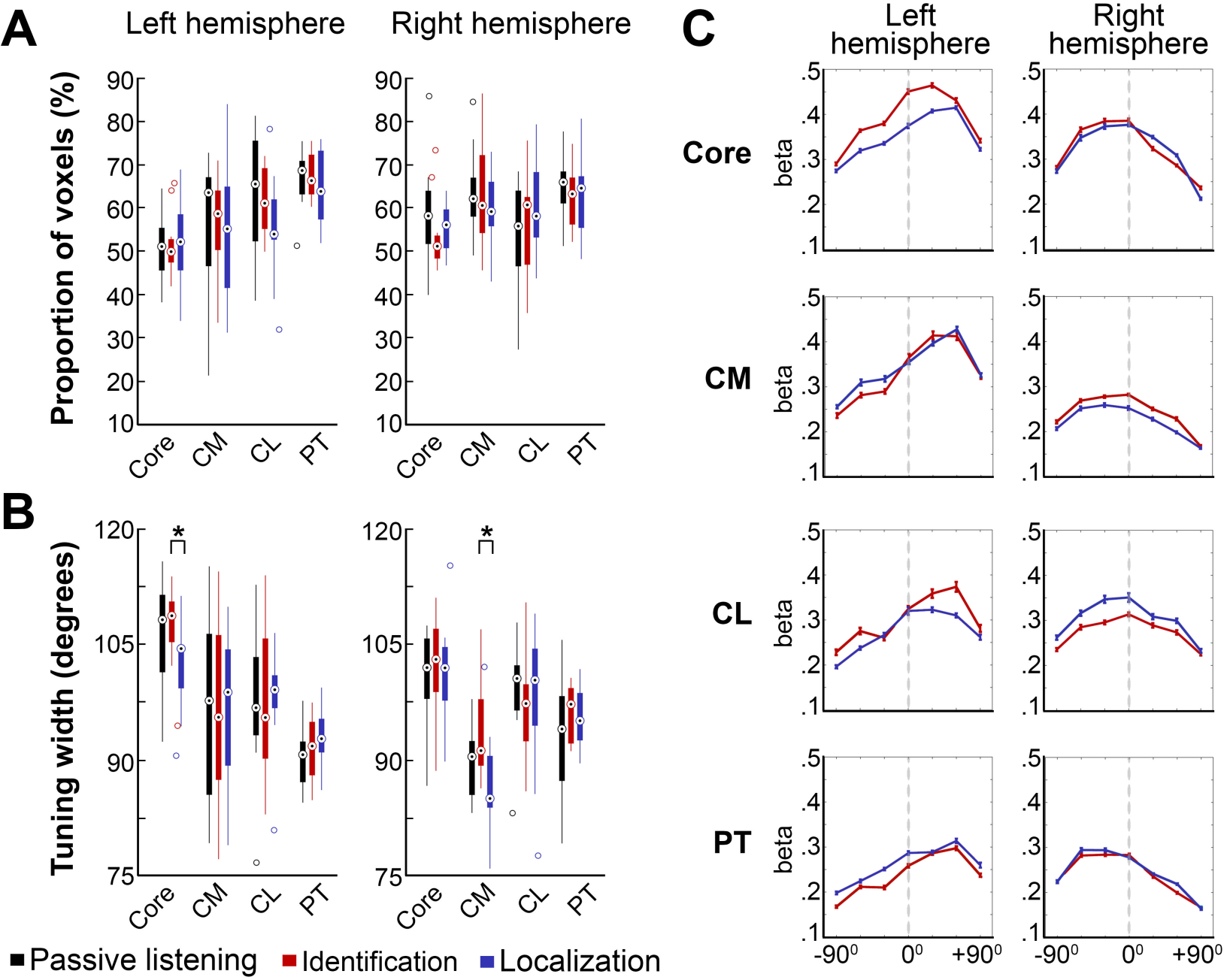
Log likelihoods for azimuth **-60⁰**

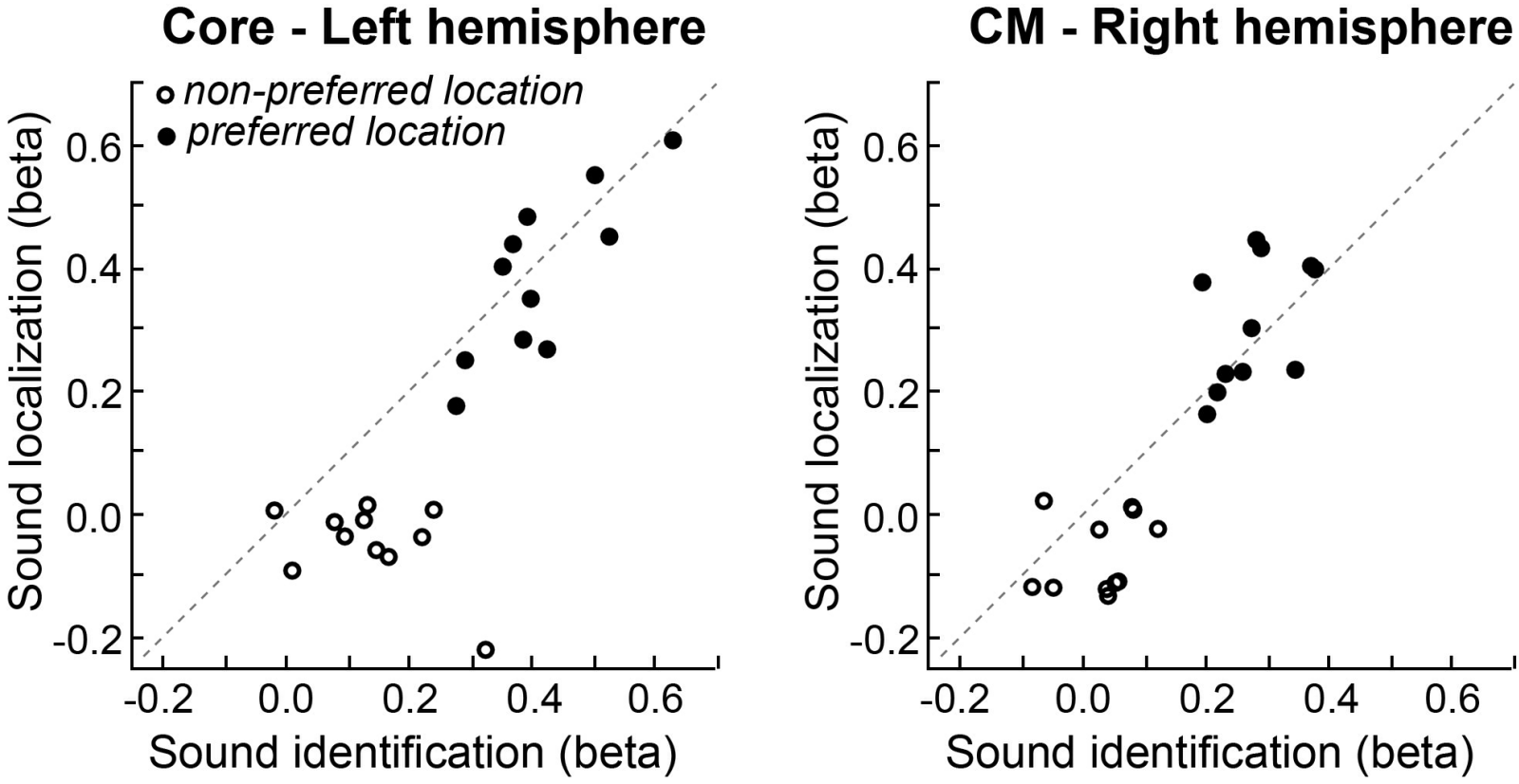
Log likelihoods for azimuth **+90⁰**

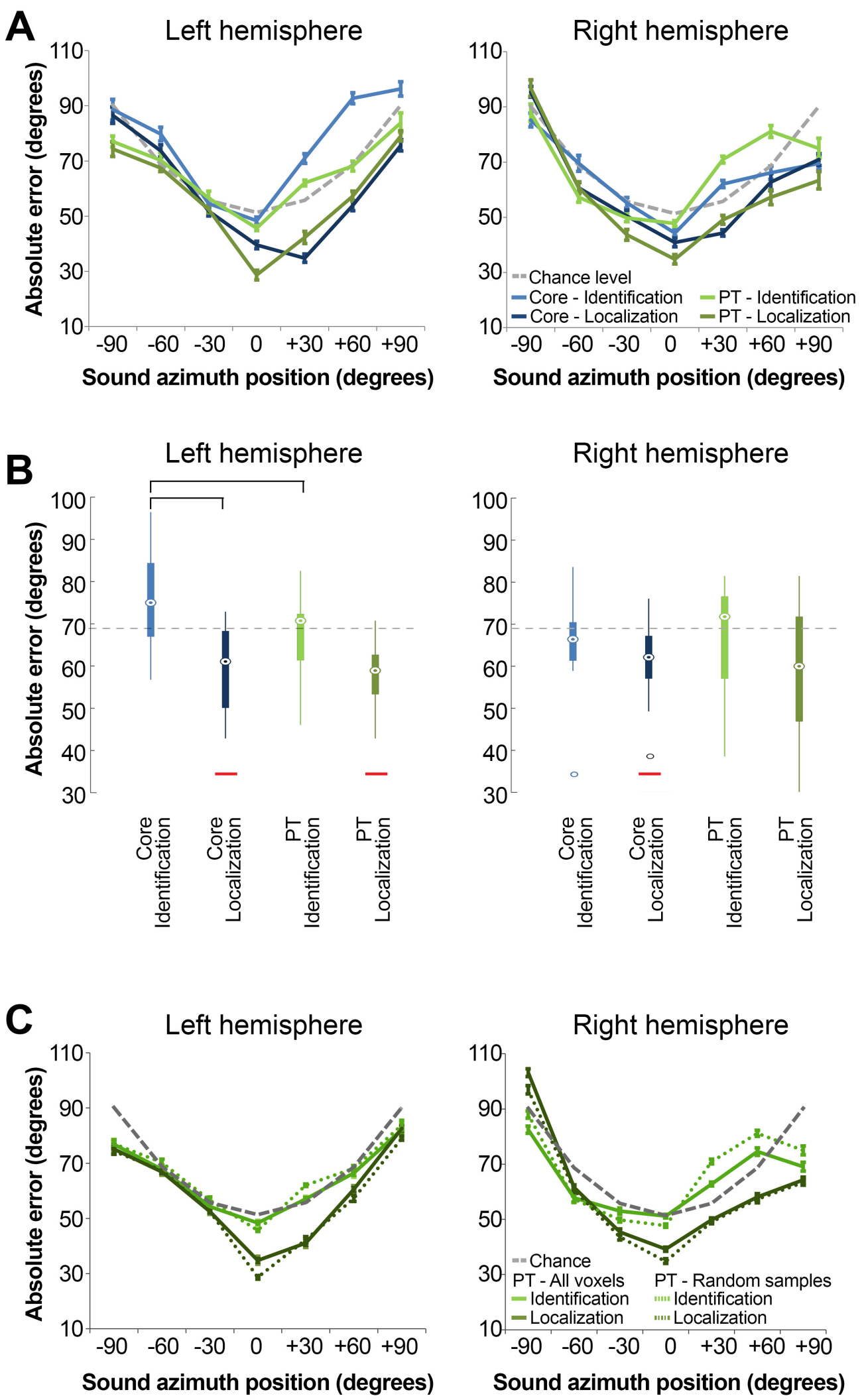
Voxel 1 Voxel 2 Voxel *N*

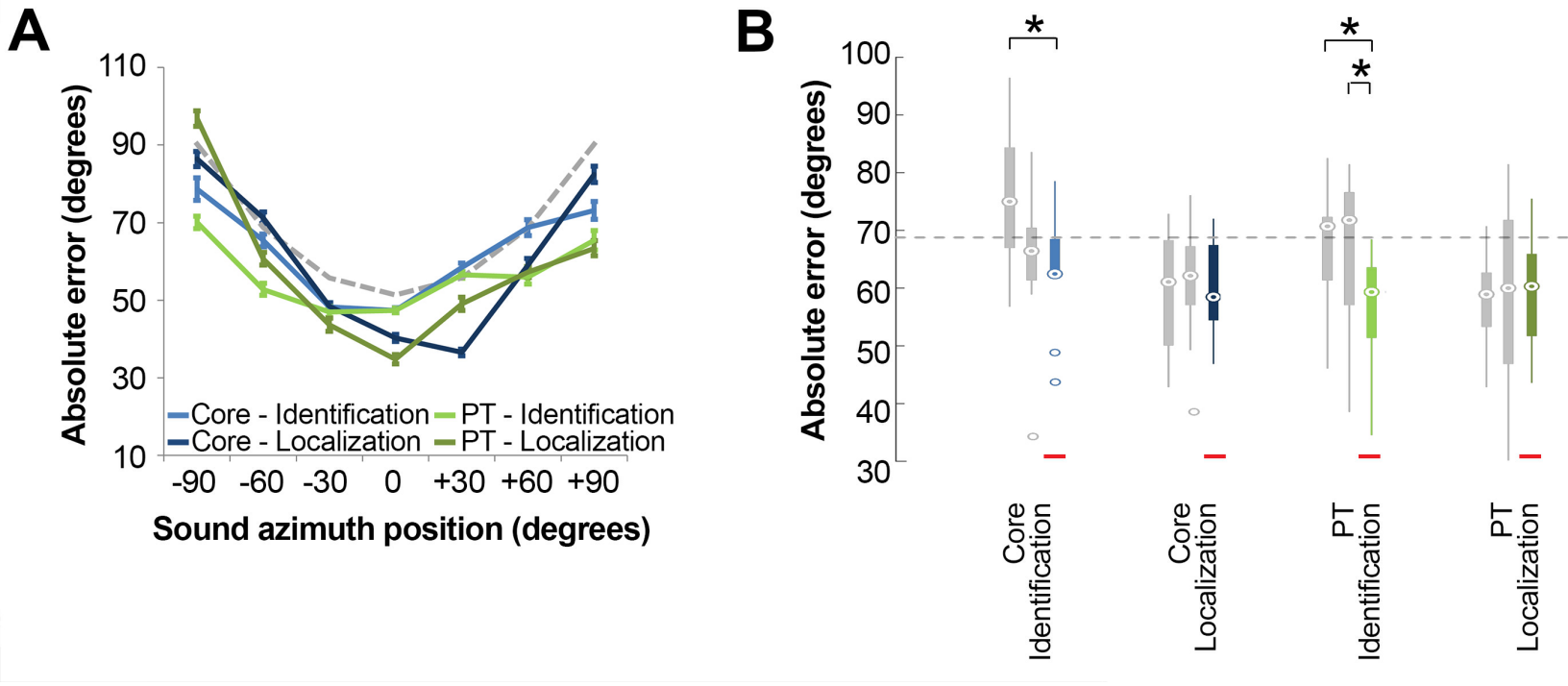

A
Planum temporale
Caudo-medial
Rostro-medial
Core
Caudo-lateral
Rostro-lateral
Dorsal
Anterior
B
LF
HF
Broad
Narrow
C
LF
HF
Broad
Narrow


A
Left hemisphere
Right hemisphere
Proportion of voxels (%)
80
70
60
50
40
30
20
Core
CM
CL
PT
Proportion (%)
50
70
B
Tuning width (degrees)
120
105
90
75
Tuning width
85⁰
105⁰


A
Left hemisphere
Right hemisphere
Proportion of voxels (%)
90
80
70
60
50
40
30
20
10
Core
CM
CL
PT
B
Tuning width (degrees)
120
105
90
75
*
Passive listening
Identification
Localization
C
Left hemisphere
Right hemisphere
Core
CM
CL
PT
beta
.5
.4
.3
.2
.1
-90⁰
0⁰
+90⁰


Core - Left hemisphere
non-preferred location
preferred location
Sound localization (beta)
Sound identification (beta)
-0.2
0.0
0.2
0.4
0.6
CM - Right hemisphere
Sound localization (beta)
Sound identification (beta)


A
Left hemisphere
Right hemisphere
Absolute error (degrees)
Sound azimuth position (degrees)
Chance level
Core - Identification
Core - Localization
PT - Identification
PT - Localization
B
Left hemisphere
Right hemisphere
Absolute error (degrees)
Core Identification
Core Localization
PT Identification
PT Localization
C
Left hemisphere
Right hemisphere
Absolute error (degrees)
Sound azimuth position (degrees)
Chance
PT - All voxels
Identification
Localization
PT - Random samples
Identification
Localization


A
Absolute error (degrees)
110
90
70
50
30
10
-90
-60
-30
0
+30
+60
+90
Sound azimuth position (degrees)
Core - Identification
PT - Identification
Core - Localization
PT - Localization
B
Absolute error (degrees)
100
90
80
70
60
50
40
30
*
*
*
Core Identification
Core Localization
PT Identification
PT Localization