

## ORIGINAL ARTICLE

# Auditory Frequency Representations in Human Somatosensory Cortex

Alexis Pérez-Bellido<sup>1,2</sup>, Kelly Anne Barnes<sup>1</sup>, Lexi E. Crommett<sup>1</sup>  
and Jeffrey M. Yau<sup>1</sup>

<sup>1</sup>Department of Neuroscience, Baylor College of Medicine, Houston, One Baylor Plaza, Houston, TX 77030, USA and <sup>2</sup>Current address: Radboud University, Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands

Address correspondence to Jeffrey M. Yau, One Baylor Plaza, T111, Houston, TX 77030, USA. Email: jeffrey.yau@bcm.edu

## Abstract

Recent studies have challenged the traditional notion of modality-dedicated cortical systems by showing that audition and touch evoke responses in the same sensory brain regions. While much of this work has focused on somatosensory responses in auditory regions, fewer studies have investigated sound responses and representations in somatosensory regions. In this functional magnetic resonance imaging (fMRI) study, we measured BOLD signal changes in participants performing an auditory frequency discrimination task and characterized activation patterns related to stimulus frequency using both univariate and multivariate analysis approaches. Outside of bilateral temporal lobe regions, we observed robust and frequency-specific responses to auditory stimulation in classically defined somatosensory areas. Moreover, using representational similarity analysis to define the relationships between multi-voxel activation patterns for all sound pairs, we found clear similarity patterns for auditory responses in the parietal lobe that correlated significantly with perceptual similarity judgments. Our results demonstrate that auditory frequency representations can be distributed over brain regions traditionally considered to be dedicated to somatosensation. The broad distribution of auditory and tactile responses over parietal and temporal regions reveals a number of candidate brain areas that could support general temporal frequency processing and mediate the extensive and robust perceptual interactions between audition and touch.

**Key words:** audio-tactile, audition, crossmodal, multisensory, touch

## Introduction

In traditional models of brain organization, sensory information is processed initially in modality-dedicated sensory cortical areas before converging in higher-order association cortex (Felleman and Van Essen 1991). Recent evidence from human neuroimaging and invasive recordings from animal models challenge this sensory processing scheme by showing that even primary sensory areas can respond to inputs from multiple sensory modalities (Ghazanfar and Schroeder 2006; Driver and Noesselt 2008). Although the occurrence of multimodal responses in sensory cortex is no longer disputed, whether the responses to “non-preferred” modalities observed in sensory

areas convey specific representations and how they relate to perception remain open questions.

Here we focus on temporal frequency processing in the human brain. We perceive temporal frequency information by both audition and touch. Temporal frequency processing is fundamental to the auditory perception of objects, speech, and music. Temporal frequency processing also underlies the tactile perception of surface textures (Manfredi et al. 2014) and indirect touch through handheld tools (Johnson 2001). Traditionally, the processing of auditory and tactile temporal frequency information has been described as segregated in separate and extensive cortical systems that reside in the temporal and parietal lobes,

respectively (Romo and Salinas 2003; Mountcastle 2005; Moerel et al. 2014). However, in a variety of behavioral contexts, audition and touch can interact in highly specific ways (Jousmaki and Hari 1998; Wilson et al. 2009, 2010; Yau et al. 2010; Occelli et al. 2011; Pannunzi et al. 2015; Crommett et al. 2017; Yau et al. 2009b), suggesting that sounds and vibrations may not be independently represented and processed. Indeed, somatosensory and auditory stimulation have been shown to co-activate a number of brain regions (Foxe et al. 2002; Kayser et al. 2005; Schurmann et al. 2006; Nordmark et al. 2012) with multimodal responses observed even in unit recordings (Fu et al. 2003). Because the vast majority of studies have focused on somatosensory responses in cortical regions traditionally considered to be dedicated to auditory function (though see, Lemus et al. 2010), much less is known about how somatosensory cortex processes auditory information. Even in cases where auditory stimulation has been shown to evoke distinct response signatures in somatosensory cortex (Liang et al. 2013), it remains unclear whether these response patterns represent specific auditory information.

In this study, we measured BOLD fMRI responses in participants performing a simple auditory frequency discrimination task using low-frequency pure-tone stimuli. Our analysis aimed to test the hypothesis that auditory representations found outside of the temporal lobes are frequency-selective and related to perception. We addressed 2 main questions using univariate and multivariate analysis techniques. First, does auditory stimulation alone modulate BOLD signals in putative somatosensory areas in a frequency-dependent manner? We tested this by characterizing auditory responses in somatosensory cortical regions that we identified on the basis of brain anatomy and independent localizer scans. By separately identifying voxels responsive to auditory, tactile, or combined audio-tactile stimulation in the localizer scans, we determined whether parietal lobe voxels exhibited selectivity for sensory modalities or responded to both auditory and tactile stimulation. This then enabled us to test whether auditory frequency representations were more likely to be supported by parietal lobe voxels that were privileged in their sensitivity to auditory inputs. Second, is there a correspondence between the perceived similarity of sounds and the sound representations carried in BOLD activation patterns found in brain regions that also respond to tactile stimulation? We addressed this question using representational similarity analyses (RSA). Crucially, because of the proximity of auditory and somatosensory areas in perisylvian cortex, we used great care to verify the localization of auditory responses in the parietal operculum in our analyses.

## Materials and Methods

### Participants

Ten participants (age range: 19–34 years; mean age: 28 years; 5 males) took part in the main fMRI experiment comprising localizer scans and scans involving auditory frequency discrimination. Six participants (age range: 19–32; mean age: 25 years; 3 males) took part in a behavioral experiment performed outside of the scanner that required subjects to rate the perceived dissimilarity of the auditory stimuli. (One of these 6 subjects also participated in the main fMRI experiment.)

Two additional control experiments were conducted (see Supplementary Material). Fourteen subjects (10 of whom participated in the main fMRI experiment plus 4 additional subjects) (age range: 19–40 years; mean age: 32 years; eight males) were

tested in a behavioral experiment that assessed the audibility of the tactile stimuli used in the main fMRI experiment. Six subjects (none of whom participated in any other experiment) (age range: 23–28 years; mean age: 26 years; one male) took part in an fMRI experiment designed to assess the dependence of brain responses on the choice of MRI-compatible headphones.

All participants were right-handed, had normal or corrected-to-normal vision, and none reported a history of auditory or somatosensory impairments. All testing procedures were performed in compliance with the policies and procedures of the Baylor College of Medicine Institutional Review Board. All participants gave their informed consent and were paid for their participation.

### Neuroimaging: Overview

The fMRI experiment involved scanning over 2 sessions on separate days (inter-session interval =  $2 \pm 1$  days). During the first session, participants became familiarized with the experimental procedures and underwent a set of anatomical and functional localizer scans designed to identify brain regions responsive to audition and touch. During the second session, participants performed a frequency discrimination task as they underwent fMRI scans. Although participants were tested using auditory and tactile stimuli, albeit in separate scans, in this report we focus our analysis on the data acquired in the scans involving auditory frequency discrimination. Behavioral and neuroimaging results for the somatosensory conditions will be presented in a separate report.

### Image Acquisition

Imaging data were acquired on a Siemens MAGNETOM Trio 3 Tesla system using a 32-channel head coil. For each participant, we obtained a structural image using a sagittal magnetization prepared rapid gradient echo (MP-RAGE) T<sub>1</sub> weighted sequence (time echo [TE] = 3.02 ms; time repetition [TR] = 2600 ms; inversion time [TI] = 900 ms; flip angle = 8°; GRAPPA factor = 2; 176 slices with 1×1×1 mm<sup>3</sup> voxels). Functional data were acquired using a 2-dimensional echo-planar imaging (EPI) sequence (TR = 2750 ms). We acquired 56 slices with 2-mm thickness in an interleaved sequence (0 mm gap, 2 × 2 mm<sup>2</sup> in-plane resolution) in an axial orientation (TE = 25 ms, flip angle = 80°; GRAPPA factor = 4). The first 3 volumes of each sequence were discarded to minimize transient effects of magnetic saturation. The same EPI pulse sequence was used for the localizer scans (5 runs; 121 samples/run) and the scans comprising the main auditory fMRI experiment (6 runs; 141 samples/run).

### Stimuli and Procedure: Functional Localizer Scans

Participants underwent 5 functional localizer runs. These scans were designed to identify brain regions that respond to auditory and/or tactile stimulation. Each localizer run consisted of a block design paradigm with 12 blocks of sensory stimulation (16.5 s) separated by intervals (11 s) during which the participants experienced no stimulation. The total duration of a functional localizer run was 354 s. Each stimulation block contained auditory, tactile or congruent audio-tactile stimulation (4 blocks/condition/run). Within a block, brief sinusoidal stimuli (duration: 0.4375 s) were delivered in alternating ascending and descending sequences that spanned 75–375 Hz (frequency steps: 75 Hz; inter-stimulus interval: 0.25 s). The frequency of the first stimulus and initial sequence direction were randomized across blocks.

Participants passively attended to the auditory and tactile stimulation patterns as they maintained visual fixation. Auditory stimulation was delivered binaurally via MRI-compatible noise-attenuating in-ear headphones (Model S14, Sensimetrics). Vibrotactile stimulation was delivered simultaneously to the distal finger pads on digits 2–5 on the right hand using a piezoelectric stimulator (CM3, Cortical Metrics).

### Stimuli and Procedure: Auditory Discrimination Scans

Participants underwent 6 scans during which they performed a perceptual task requiring them to attend to auditory stimulus frequency. Auditory stimuli tested in the main fMRI experiment comprised 5 pure tones of 75, 130, 195, 270, and 355 Hz (duration: 0.500 s; rise/fall time: 0.01 s; sample rate: 44.1 kHz). These stimuli were selected to ensure that none of the tested frequencies were harmonically related. Because we were interested in characterizing response pattern differences related to variations in stimulus frequency, we presented stimuli at amplitudes that were equated for perceived intensity (range: 45.3–57.9 dB SL) as established in preliminary experiments. To ensure that stimulus amplitude did not systematically vary with stimulus frequency, we introduced a jitter ( $\pm 5\%$ ) to the loudness-matched amplitude levels on every stimulus presentation. The auditory stimuli were generated in Matlab (2011b; Mathworks) running on a Macbook and were delivered through the MRI-compatible headphones after amplification (PCA1, Pyle).

We used an event-related design in the main auditory discrimination experiment (duration/scan: 406 s). Each event (2.75 s) consisted of the sequential presentation of 4 stimuli (stimulus duration: 0.5 s; inter-stimulus interval: 0.25 s). Events fell into one of 2 categories (Fig. 1A). In a standard event, the 4 stimuli were of the same frequency. In an oddball event, one of the stimuli (either the second or third stimulus) differed in frequency from the others in the sequence. Each scan consisted of 35 events comprising 30 standard events (6 repetitions of each test frequency) and 5 oddball events. Event times in each scan (inter-event intervals ranged from 5.5 to 13.75 s) were pre-defined using Optseq2 (<http://surfer.nmr.mgh.harvard.edu/optseq>) and the scan order was counter-balanced

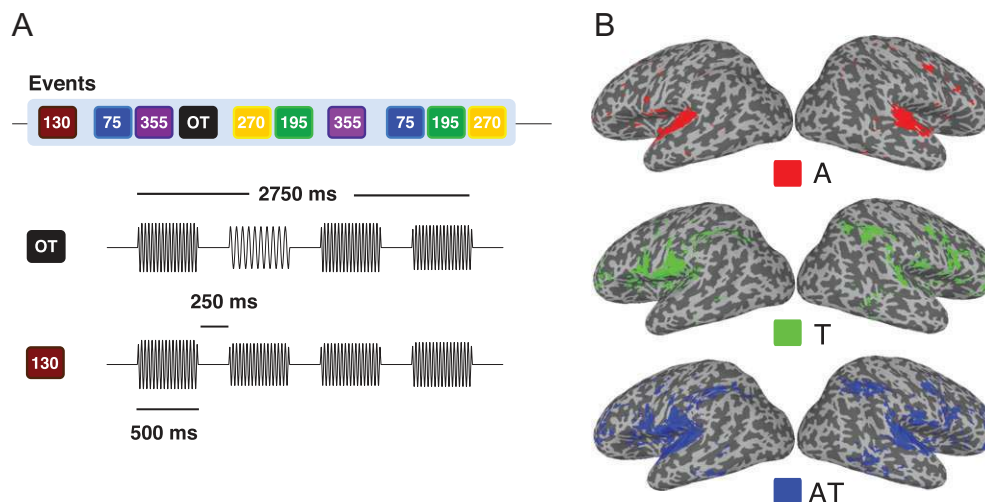
across participants. Participants were instructed to indicate the occurrence of oddball events by button press using a response device held in their left hand. We required responses to occur within 2 s after the presentation of the test stimuli. A visual fixation cross, presented in black throughout the experiment, temporarily turned gray when a button press was registered. At the conclusion of each scan, participants received performance feedback (% correct). Participants were familiarized with the stimuli, response device, and task before scanning.

### Behavioral Experiment: Frequency Similarity Ratings

In separate experiments performed outside of the scanner, participants performed a similarity-rating task using the stimuli tested in the main fMRI experiment. The experiment consisted of 2 blocks each comprising 50 trials. Each trial contained 2 stimulus intervals, separated by 0.5 s, during which 4-tone sequences were delivered via in-ear headphones (Sennheiser CX200). Participants rated the perceived similarity of the frequency tone sequences presented in the 2 trial intervals using a scale ranging from 1 (identical) to 5 (highly dissimilar). The 4-tone sequences comprising each interval were sampled randomly from the standard events tested in the main fMRI experiment. There were a total of 15 possible sequence pairs: 5 pairs (4 repetitions each) contained sequences that were matched in frequency and 10 pairs (8 repetitions each) contained sequences that differed in frequency. Participants had 4.5 s to respond on each trial. Participants were familiarized with the stimuli and test procedures in a training block immediately before the start of the 2 test blocks.

### Data Analysis: Overview

We first identified the brain regions that responded to auditory, tactile, or congruent audio-tactile stimulation in the functional localizer scans. Regions identified as responsive to the different unimodal and audio-tactile conditions in the localizer data comprised separate masks that constrained the data space in the analyses of the frequency discrimination scans. Using separate functional masks enabled us to test whether the analysis of the auditory discrimination scans yielded different results



**Figure 1.** Auditory discrimination scan design and identification of sensory regions. (A) Auditory stimuli (75, 130, 195, 270, 355 Hz) were presented during pure tone (standard) trials and oddball trials (OT) in an event-related design. Events consisted of 4-tone sequences. On OT events, 3 of the 4 tones were matched in frequency. On standard events, all 4 tones were matched in frequency. Participants were instructed to report the occurrence of OT events during the auditory discrimination scans. (B) Results from the block-design functional localizer scans (Materials and Methods) depicting regions that responded to auditory (A; red), tactile (T; green), and concurrent audio-tactile (AT; blue) stimulation ( $P < 0.05$ , uncorrected).

when constrained to regions responsive to the unimodal or audio-tactile conditions in the functional localizer scans. In univariate analyses of the auditory discrimination scans performed in surface space, we conducted group-level statistical tests of parameters estimated using general linear models (GLM) fitted to the time series of each element of the surface mesh (i.e., node). These analyses identified the brain regions where BOLD signals at the surface node level varied according to auditory stimulation. In multivariate analyses, we conducted representational similarity analyses (RSA; Nili et al. 2014) using auditory and somatosensory regions-of-interest (ROI) defined using a combination of functional and anatomical masks. These analyses characterized relationships among the multi-voxel activation patterns associated with auditory stimulation presented at different frequencies as well as their relationship to perception. The univariate and multivariate analyses are described in detail in the following sections.

### fMRI Data Pre-processing

We used AFNI software (Cox 1996; AFNI\_16.1.19) to perform data pre-processing and univariate analyses. Surface models were created with FreeSurfer (Fischl et al. 1999; freesurfer-Darwin-lion-stable-pub-v5.3.0) and visualized in SUMA (Saad and Reynolds 2012). Functional datasets were despiked after correcting for slice timing and head motion. Volumes with head motion exceeding 0.3 mm/TR or containing a substantial number outlier voxels (more than 5% of total voxels) defined according to time series statistics (3dToutcount function) were excluded from analysis. Data from the localizer scan and the frequency discrimination scans that were included in whole-brain univariate analyses were projected on standard surface meshes (Saad and Reynolds 2012; linear depth: 64) and spatially smoothed (4-mm FWHM 2D Gaussian kernel) for group-level analysis. No warping or smoothing was applied to the volumetric data from the frequency discrimination scans that were included in the RSA. All data were expressed in percent signal change with respect to the mean signal in each scan.

### fMRI Data Analyses: Localizer Scans

The functional localizer data were analyzed in surface space. Data from 2 runs in one subject and 3 runs in another subject were excluded from analyses because of excessive image artifacts. Localizer scans were modeled using GLMs that included 3 regressors of interest corresponding to the blocks of auditory, tactile, and audio-tactile stimulation convolved with gamma-variate functions. Head motion parameters and drift parameters (linear, quadratic, and cubic) were included as nuisance regressors. In group-level analyses, we identified nodes that exhibited positive BOLD signal changes in each unimodal or audio-tactile stimulation block. Because we intended to use these response patterns to constrain the univariate analyses of the auditory discrimination scans, we used a liberal threshold  $P < 0.05$  (uncorrected) to generate separate masks for the auditory, tactile, and audio-tactile conditions (Fig. 1B).

### fMRI Data Analyses: Auditory Discrimination Scans

For all analyses, the frequency discrimination scans were first modeled using GLMs that included 6 regressors of interest comprising the 5 frequency conditions and 1 oddball event condition, along with nuisance regressors. Gamma-variate functions were used to model stimulus responses in all analyses. In a

separate validation procedure, we analyzed the data with finite impulse response deconvolution in order to estimate response time courses for each stimulus condition (8 time points; 0–19.25 s post-stimulus onset).

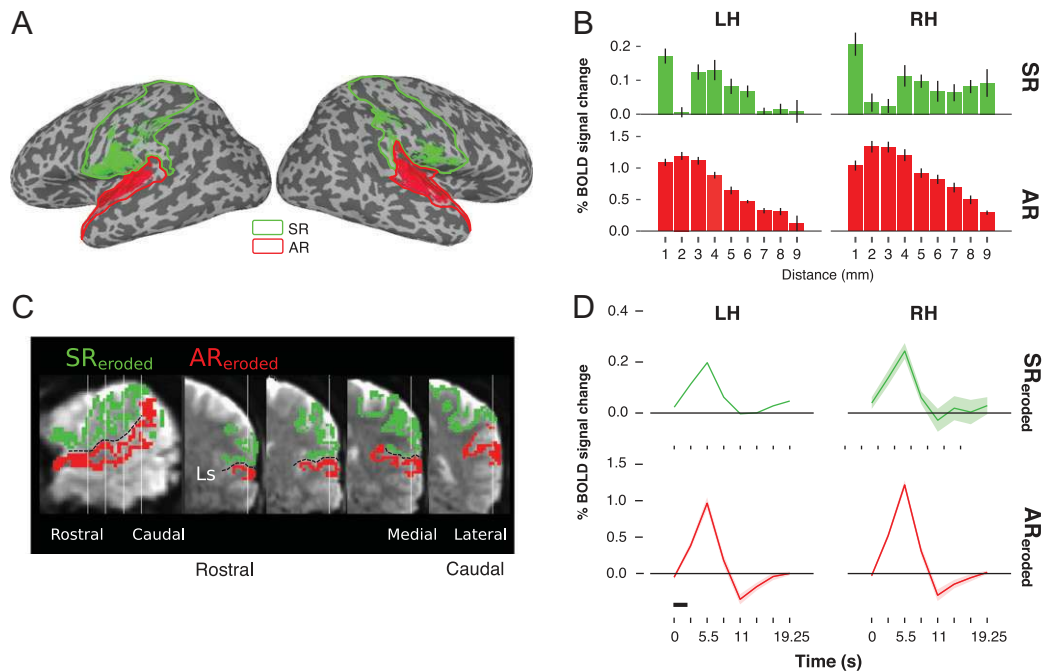
In group-level univariate analyses performed in surface space, we identified nodes exhibiting significant responses (i.e., positive BOLD signal changes relative to baseline) to auditory stimulation, regardless of stimulus frequency, using paired *t*-tests. We identified nodes exhibiting significant frequency-dependent response modulation using a one way repeated-measures ANOVA with subject as random factor and frequency (75, 130, 195, 270, and 355 Hz) as a fixed factor. Separate *t*-tests and ANOVAs were conducted using the auditory, tactile, and audio-tactile masks generated from the localizer scans (using AFNI functions 3dtttest++ and 3dANOVA2, respectively). All group-level univariate test results were statistically thresholded at a false discovery rate (FDR) corrected  $q < 0.05$ . We used the “SurfClust” AFNI function to summarize the clustered activation patterns identified as significant after FDR correction (Supplementary Material).

### ROI Mask Generation

Due to the spatial proximity of somatosensory cortical regions in the parietal operculum and auditory cortical regions in the temporal lobe, attribution of fMRI BOLD responses to these perisylvian areas can be particularly challenging (Ozcan et al. 2005). Accordingly, we defined a set of subject-specific ROIs for our analysis of the auditory discrimination scans so that we could account for voxels located at the borders separating the parietal and temporal lobes (see below). The logic of our 3-part approach is summarized here and then detailed in the paragraphs below: First, we defined a set of general parietal lobe and temporal lobe anatomical ROI masks for each hemisphere separately. Note that we used general ROI definitions simply to facilitate comparisons between somatosensory regions in the parietal lobe and auditory regions in the temporal lobe. Second, we evaluated the spatial profile of auditory responses at the boundary between the general parietal and temporal ROI masks. This evaluation allowed us to identify voxels in perisylvian cortex that may be particularly vulnerable to partial volume effects. Finally, to better ensure that partial volume effects did not confound our ROI-based analyses, we generated conservative ROI masks that excluded vulnerable voxels and used these “eroded ROI” masks in the data analysis. We now explain the details of each of these steps.

First, in each hemisphere, we defined a general somatosensory ROI and a general auditory ROI (Fig. 2A; outlined boundaries). These general ROI masks were constructed by combining anatomical parcellations from the Destrieux atlas (Destrieux et al. 2010), which were derived separately for each participant based on their unique brain topologies. The general ROI spanning putative somatosensory regions (SR) comprised the central sulcus, postcentral gyrus, postcentral sulcus, supramarginal gyrus, subcentral gyrus, and subcentral sulcus. The general ROI spanning putative auditory regions (AR) comprised the posterior segment of the lateral fissure, planum temporale, transverse temporal sulcus, superior temporal gyrus, transverse temporal gyrus, lateral aspect of the superior temporal gyrus and the inferior segment of the circular sulcus of the insula.

Second, we used the general SR and AR ROI masks to evaluate the spatial distribution of the auditory responses measured in the discrimination scans. The purpose of this preliminary analysis was to ensure that auditory responses found outside of AR were not confined solely to voxels found at the border between the parietal and temporal lobes. This evaluation



**Figure 2.** Defining somatosensory and auditory regions of interest. (A) Groupings of anatomical masks (Materials and Methods) defined general regions of interest corresponding to somatosensory regions (SR; green) and auditory regions (AR; red) in the left and right hemispheres (LH and RH, respectively). The extended regions defined by the groupings of the anatomical masks are indicated by the colored boundaries. Filled-in portions within the colored boundaries indicate regions active during any of the localizer conditions. (B) BOLD signal changes associated with auditory stimulation in SR and AR in each hemisphere. Bar plots indicate the mean response over all frequency conditions averaged across SR (AR) voxels sorted according to their proximity to the nearest voxel in AR (SR). Error bars indicate response variance across frequency conditions averaged over voxels in each distance bin. (C) Sagittal and coronal slices show example EPI data (averaged over one run) from a single subject. Colored overlays depict the voxels residing within  $SR_{eroded}$  and  $AR_{eroded}$  (excluding voxels falling within 2 mm of the border between SR and AR) that were potentially included in the univariate analyses depending on the localizer mask condition. The dashed line indicates the lateral sulcus (Sylvian fissure). (D) Temporal response profiles show group-averaged BOLD signal time courses for the  $SR_{eroded}$  and  $AR_{eroded}$  in the left and right hemispheres. The time courses represent the response profile associated with each voxel's preferred frequency averaged over all of the voxels in the ROI. The horizontal black line indicates the time of stimulus presentation. Shaded areas represent SEM.

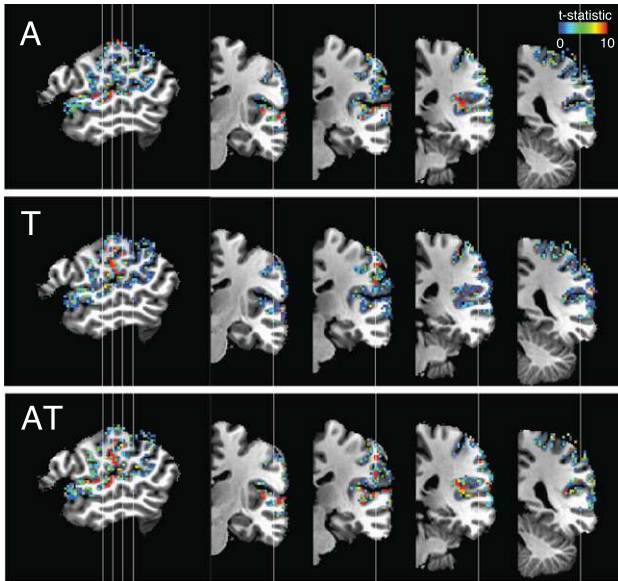
addresses the concern that partial volume effects or other confounding spatial artifacts might trivially explain auditory responses in SR. We restricted this evaluation to all of the active nodes identified from the localizer scans (i.e., union of the nodes responsive in the auditory, tactile, and audio-tactile conditions) which fell within SR and AR (Fig. 2A, filled nodes) and quantified the BOLD signal changes associated with auditory stimulation in the discrimination scans. In native volume space, we calculated the average BOLD signal change as a function of the Euclidean distance from the ROI borders for each subject. Taking SR as an example, we averaged the mean response over all frequencies across all voxels that fell within a particular distance (in 1-mm steps) defined with respect to the nearest voxel within AR (red boundary in Fig. 2A). We then averaged these spatial response profiles over subjects, and we repeated this procedure for SR and AR in both hemispheres. Visual inspection of the group-averaged spatial response profiles (Fig. 2B) revealed robust responses over a large distance range. The clear inflections in the response profiles occurring within 2 mm of the ROI borders are consistent with some partial volume effects or signal blurring at the ROI boundaries even without explicit spatial smoothing of the data during pre-processing.

Finally, to account for these potential partial volume effects or signal blurring in our ROI-based multivariate analyses (see below), we constrained the SR and AR masks by excluding voxels falling within 2 mm of the ROI borders. This ROI-eroding procedure yielded smaller and more conservative ROIs (see  $SR_{eroded}$  and  $AR_{eroded}$  in Fig. 2C). Despite the exclusion of boundary voxels,

robust and consistent auditory responses were evident in  $SR_{eroded}$  and  $AR_{eroded}$  in both hemispheres (Fig. 2D). While response magnitudes were generally larger in auditory regions compared to somatosensory regions, response time courses in all ROIs followed a stereotypical temporal profile with signals peaking approximately 5.5 s after stimulus onset before decaying back to baseline (Rosen et al. 1998). Because they provide an objective way to address the potential impacts of partial volume effects in perisylvian cortex, we used  $SR_{eroded}$  and  $AR_{eroded}$  in ROI-based extensions of the group-level univariate analyses. In these analyses, ROI-averaged responses were first calculated for individual subjects, from active voxels contained within  $SR_{eroded}$  and  $AR_{eroded}$  (identified as active during the localizer scans with an uncorrected  $P < 0.05$  threshold), before averaging at the group level.

### Representational Similarity Analysis (RSA)

We used RSA (Nili et al. 2014; <http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/>) to characterize the similarity of multi-voxel response patterns to the different auditory stimuli tested in the discrimination scans. These analyses were performed in native volume space and independent analyses were conducted for voxels identified from the auditory, tactile, and audio-tactile localizer conditions. Note that voxel selection using the localizer masks for RSA, unlike the ROI-based group-level univariate analyses, was performed at a single subject level. For each subject separately (Fig. 3), RSA was performed using all of the voxels in a given localizer mask (e.g., the auditory condition) that



**Figure 3.** Localizer scan results for an example subject. Colors indicate t-statistic values for voxels falling within  $SR_{eroded}$  and  $AR_{eroded}$ . Separate t-maps are shown for the auditory (A), tactile (T), and audio-tactile (AT) blocks in the localizer scan. Subject-specific t-maps were used for voxel selection in the representational similarity analysis (Materials and Methods).

exceed a threshold of  $t > 0$  and that fell within the anatomical mask of interest (e.g.,  $SR_{eroded}$ ). Dissimilarity metrics were then aggregated over subjects for group-level statistical tests. Performing RSA using the different modality-defined ROI masks allowed us to test the hypothesis that auditory frequency representations are more robust in voxels that responded preferentially to auditory stimulation during the localizer scans. Over a series of analyses, we characterized the representational similarity of the auditory stimuli in a number of ROIs. We first performed RSA on  $SR_{eroded}$  and  $AR_{eroded}$  combined across hemispheres to characterize representational geometries over the largest and most general definitions of somatosensory and auditory regions. Second, we performed RSA on  $SR_{eroded}$  and  $AR_{eroded}$  separately in each hemisphere to explore laterality effects. Lastly, for a first approximation of the representational spaces occupied by sounds across different putative levels of the somatosensory cortical hierarchy, we performed RSA on voxels falling within Destrieux atlas parcellations of the parietal lobe (central sulcus, postcentral gyrus, postcentral sulcus, supramarginal gyrus, and subcentral gyrus and sulcus). These parcels, in order, loosely correspond to primary somatosensory cortex and higher-order somatosensory regions. In this analysis, the data for each ROI were collapsed over the 2 hemispheres because we did not find evidence for hemispheric differences in the analyses exploring laterality effects (see below).

For each ROI, we combined the response amplitudes (i.e., beta coefficients from the first-level GLM) over all of the included ROI voxels to form separate activity vectors for each auditory frequency. Each vector can be understood as the distributed neural representation of a particular auditory stimulus across the included ROI voxels. To estimate the similarity between the response patterns corresponding to the 5 tested auditory frequencies, we calculated the Pearson correlation between each pairwise combination of activity vectors. For each ROI separately, we generated an  $m \times n$  cortical dissimilarity matrix (DM) in which each entry contained a distance metric equal to 1 minus the

correlation between the  $m$ th and  $n$ th activity vectors. Thus, distance values of 0 (such as those in the diagonal entries) indicate identical activity patterns and large distance values indicate highly dissimilar activity patterns. For each subject separately, we quantified the relationship between cortical dissimilarity values and absolute frequency differences between the stimuli by calculating Spearman correlations and we tested the significance of this relationship in each ROI at the group level using one-sample non-parametric Wilcoxon signed-rank tests. For each analysis, all of the effects reported as achieving statistical significance included adjustments for the number of ROI using FDR correction ( $q < 0.05$ ); however, the uncorrected  $P$ -values for each analysis are reported in the text.

We also generated dissimilarity matrices based on perceptual dissimilarity ratings of the auditory tones acquired in the offline behavioral experiments. Ratings for each pair of auditory tones were averaged across participants to generate a group perceptual DM. The distance values in the group perceptual DM were normalized such that the distances ranged from 0 (most similar) to 1 (least similar). We quantified the relationship between perceptual dissimilarity values and absolute frequency differences between the stimuli by calculating Spearman correlations and we tested the significance of this relationship at the group level using one-sample non-parametric Wilcoxon signed-rank tests.

To relate neural and perceptual representations of the auditory frequencies, we compared the cortical DM calculated for each ROI to the group perceptual DM. We computed the Spearman correlation between the perceptual DM and each participant's cortical DMs. A high correlation in these analyses would indicate that the representational geometry of the auditory stimuli in the ROI corresponds to the geometry of the stimuli in perceptual space. At the group-level, the statistical significance of the correlations between perceptual and cortical DMs was determined using one-sample non-parametric Wilcoxon signed-rank tests (FDR corrected). We estimated 95% confidence intervals for the Spearman correlations using a non-parametric bootstrapping procedure (1000 permutations).

## Results

### Behavioral Results: Auditory Discrimination Scans

Participants were highly accurate in detecting oddball events (mean  $d' = 5.89$ ; 95% CI [4.84, 6.93]). These high performance levels demonstrate that participants were highly sensitive to auditory frequency differences and reliably attended to the frequency of the auditory stimulation during the auditory discrimination scans.

### Univariate Analysis Results: Auditory Discrimination Scans

We characterized the relationship between BOLD responses and auditory stimulation in the discrimination scans using node-wise univariate analyses. Independent tests were performed using the auditory, tactile, and audio-tactile masks derived separately from the localizer scans. Note that, although there was substantial overlap between the activation patterns associated with the unimodal and bimodal blocks in the localizer scans (Dice coefficients: AvsT = 0.17; AvsAT = 0.46; TvsAT = 0.42), there were clear differences in these patterns (Fig. 1B). With the auditory localizer mask, we found that auditory stimulation during the discrimination scans, collapsed over frequency conditions, was associated with significant signal modulation (FDR  $q < 0.05$ ) in 37 and 54 clusters spanning the

left and right hemispheres, respectively (Fig. 4A, top). Outside of extensive bilateral superior temporal lobe regions, significant auditory responses were also observed in the parietal operculum and frontal cortex. With the tactile localizer mask, auditory stimulation was associated with significant responses in 42 (left hemisphere) and 74 (right hemisphere) clusters extending largely over frontal and parietal regions, but also including bilateral portions of the temporal lobe (Fig. 4A, middle). With the audio-tactile localizer mask, significant auditory activations were observed in 42 (left hemisphere) and 110 (right hemisphere) clusters that overlapped substantially with the activations found using the unimodal localizer masks (Fig. 4A, bottom). Supplementary Table 1 lists the locations and spatial extents of the significant clusters identified in each analysis with sizes exceeding 40 surface nodes. In order to account for partial volume effects that could have produced some of the auditory responses in parietal cortex, we further restricted the univariate analyses using eroded ROI (Materials and Methods). Despite excluding data recorded near the parieto-temporal border, significant auditory responses were observed in both  $AR_{eroded}$  and  $SR_{eroded}$ . Across ROI and hemispheres (Fig. 4B), larger responses were found using the auditory mask compared to the tactile and audio-tactile masks. Collectively, these analyses indicate that auditory stimulation during the discrimination scans associated with significant BOLD signal changes in distributed brain regions including, but not restricted to, classically defined auditory areas.

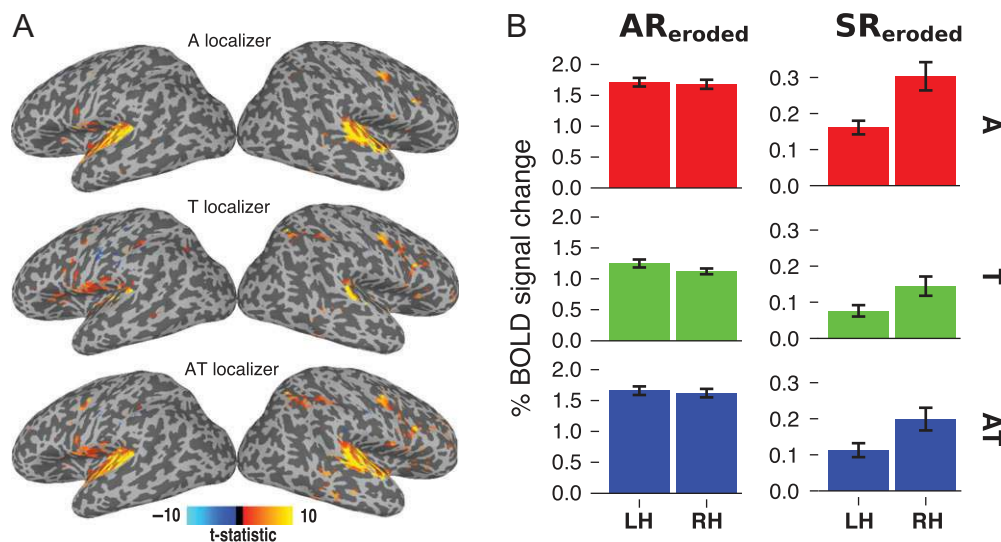
We next tested whether the magnitude of BOLD signal changes during the discrimination scans depended on auditory stimulus frequency. With the auditory localizer mask, a one way repeated-measures ANOVA revealed a significant main effect of frequency (FDR  $q < 0.05$ ) in 21 (left hemisphere) and 27 (right hemisphere) clusters. Most of these clusters resided in extensive bilateral superior temporal lobe regions, but small clusters were also located in the left frontoparietal operculum and right frontal cortex (Fig. 5A, top). In contrast, substantially fewer and smaller regions exhibited a significant main effect of frequency in the

analysis using the tactile localizer mask (Fig. 5A, middle; left hemisphere: 6 clusters, right hemisphere: 13 clusters). With the audio-tactile localizer mask, significant frequency-dependent effects were observed in both hemispheres in the upper and lower banks of the Sylvian fissure (Fig. 5A, bottom; left hemisphere: 27 clusters, right hemisphere: 32 clusters). The majority of these clusters fell within the temporal lobe, but a number of clusters were also located in parietal and frontal regions. Supplementary Table 2 lists the locations and spatial extents of the clusters exhibiting significant main effects of frequency attained with each mask. Even after accounting for vulnerability to partial volume effects, robust and significant responses to the different auditory stimulus frequencies were observed in the somatosensory and auditory ROIs (Fig. 5B). In general, larger responses were found using the auditory mask compared to the tactile and audio-tactile masks. Furthermore, response profiles over the stimulus frequencies were similar across hemispheres and localizer masks (Fig. 5B), although these profiles differed slightly between  $AR_{eroded}$  and  $SR_{eroded}$ . Specifically, in both regions qualitatively larger responses were centered on 200 Hz, but the bimodally distributed responses in  $SR_{eroded}$  also contained a peak at 75 Hz, the lowest tested frequency. These univariate results reveal that frequency-dependent BOLD signal changes that were associated with auditory stimulation occurred within and outside of classically defined auditory regions. The differences in the results acquired with the tactile localizer masks compared to the auditory and audio-tactile masks imply that brain regions preferentially responsive to auditory stimulation in the localizer scans were more likely to exhibit significant response modulation related to auditory frequency.

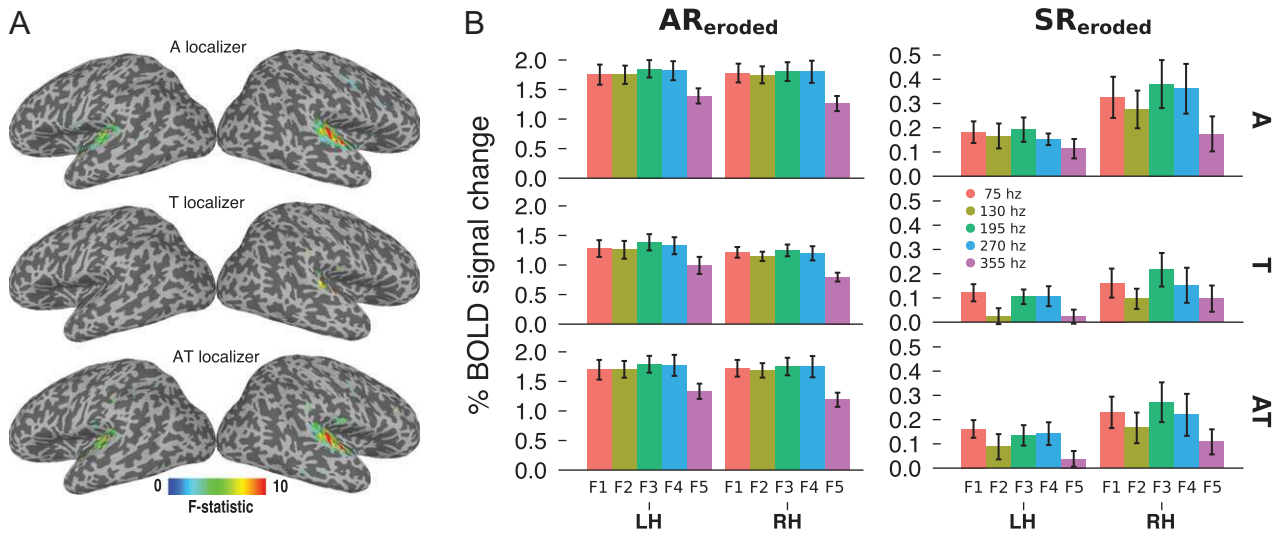
## Representational Similarity Analysis Results

### RSA on Bilateral ROI

Are there systematic relationships between the multivariate response patterns associated with each auditory frequency in the parietal and temporal lobes? If so, do these relationships



**Figure 4.** Regions generally active during the discrimination scans irrespective of sound frequency. (A) Activations indicate regions that exhibited significant responses to auditory stimulation relative to baseline irrespective of stimulus frequency in group-level tests. Activations are shown at a threshold of  $q < 0.05$  (constrained by functional localizer mask, FDR corrected) for the auditory (A), tactile (T), and audio-tactile (AT) localizer masks separately. (B) For each localizer mask separately (A, T, and AT), bar plots indicate the auditory response magnitude averaged over frequency conditions in the somatosensory ( $SR_{eroded}$ ) and auditory ( $AR_{eroded}$ ) ROI in the left and right hemispheres (LH and RH, respectively). Error bars indicate SEM. Note that the data plotted for the somatosensory and auditory ROIs are shown on different scales.



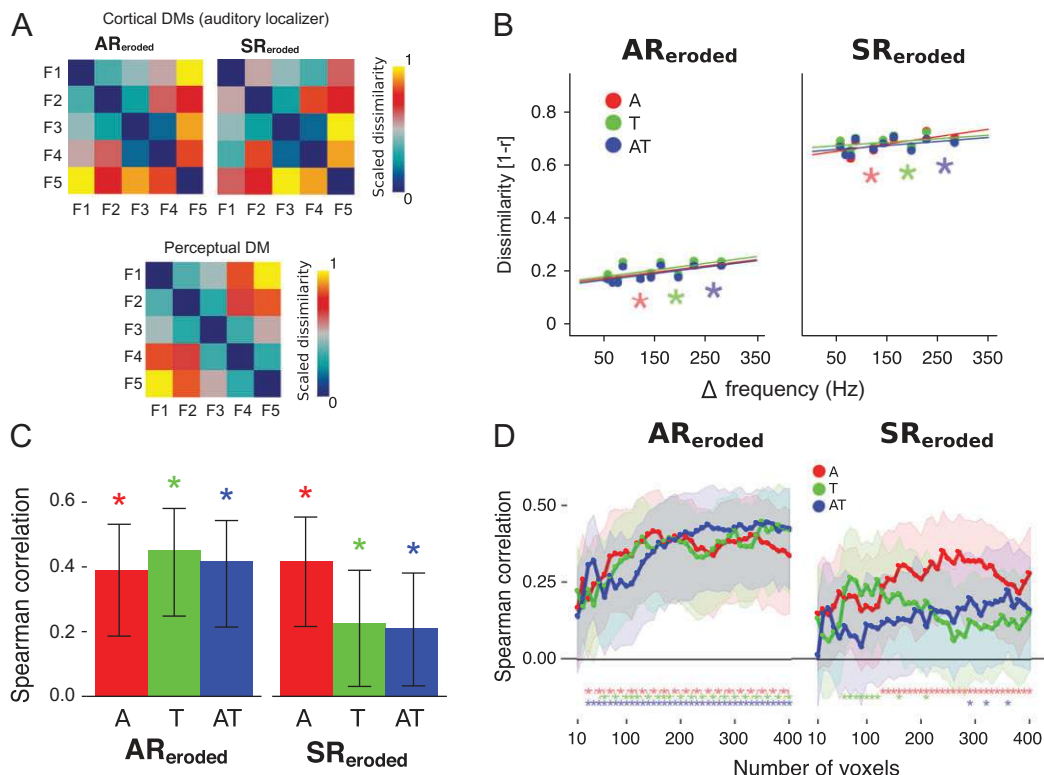
**Figure 5.** Regions exhibiting frequency-dependent response modulation during the discrimination scans. (A) Activations indicate regions in which auditory response magnitudes significantly varied according to stimulus frequency in group-level tests. Activations are shown at a threshold of  $q < 0.05$  (constrained by functional localizer mask, FDR corrected) for the auditory (A), tactile (T), and audio-tactile (AT) localizer masks separately. (B) For each localizer mask separately (A, T, and AT), bar plots indicate the average auditory response magnitude for each frequency condition (75, 130, 195, 270, and 355 Hz) in the somatosensory (SR<sub>eroded</sub>) and auditory (AR<sub>eroded</sub>) ROI in the left and right hemispheres (LH and RH, respectively). Error bars indicate S.E.M. Note that the data plotted for the somatosensory and auditory ROIs are shown on different scales.

reflect the manner by which the auditory stimuli are perceived? Furthermore, do these relationships differ depending on the auditory, tactile, and audio-tactile localizer masks? To address these questions, we explored the representational geometry of the auditory stimuli using RSA. We first focused on multivariate response patterns in bilateral AR<sub>eroded</sub> and SR<sub>eroded</sub> (Materials and Methods). With the auditory localizer mask, the spatial activation patterns in the bilateral auditory and somatosensory ROIs were distinctly related to stimulus frequency (Fig. 6): stimulus pairs that were more similar in frequency tended to have more similar spatial response patterns. For example, the group-averaged cortical DM for AR<sub>eroded</sub> (Fig. 6A) shows that the response patterns associated with the 75-Hz stimulus tended to be highly dissimilar to the patterns associated with the 355-Hz stimulus (i.e., characterized by larger distance). Conversely, stimulus pairs consisting of relatively small differences in frequency tended to have more similar activation patterns. These systematic tendencies were also evident in the cortical DM for the bilateral SR<sub>eroded</sub>. For each localizer mask separately, we quantified the strength and significance of these dissimilarity patterns in bilateral AR<sub>eroded</sub> and SR<sub>eroded</sub>. In the auditory ROI (Fig. 6B), the positive relationship between frequency distance and pattern dissimilarity was highly significant regardless of localizer condition (auditory mask: mean  $r = 0.55$ ,  $W = 55$ ,  $P < 0.001$ ; tactile mask: mean  $r = 0.56$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile mask: mean  $r = 0.57$ ;  $W = 55$ ;  $P < 0.001$ ). In the somatosensory ROI (Fig. 6B), a stronger relationship between frequency distance and pattern dissimilarity was observed with the auditory mask, but significant positive relationships were evident with all masks (auditory mask: mean  $r = 0.40$ ,  $W = 54$ ,  $P < 0.005$ ; tactile mask: mean  $r = 0.24$ ;  $W = 45$ ;  $P < 0.05$ ; audio-tactile mask: mean  $r = 0.25$ ;  $W = 45$ ;  $P < 0.05$ ). The higher overall dissimilarity values in SR<sub>eroded</sub> as compared to AR<sub>eroded</sub> are due to the fact that auditory activation patterns in the somatosensory ROI tended to be noisier and less robust than activation patterns in the auditory ROI. Despite differences in overall dissimilarity metrics, auditory frequency information is represented in the

spatial activity patterns in both parietal and temporal lobe regions; however, the larger Spearman correlations in AR<sub>eroded</sub> imply that activation patterns in temporal regions are qualitatively more reliable for discriminating auditory stimulus frequency.

As a first approximation to understanding how the frequency representations in each ROI related to the perception of the auditory tones, we quantified the perceived dissimilarity of the sound stimuli by generating a perceptual DM from the similarity ratings recorded in separate behavioral experiments (Materials and Methods) and then compared the perceptual DM to the cortical DMs calculated for each bilateral ROI. The group-averaged perceptual DM (Fig. 6A) shows that participants rated the auditory stimulus pairs containing more similar frequencies as more similar perceptually. This relationship between perceptual similarity ratings and absolute differences in stimulus frequency was positive and significant (mean  $r = 0.83$ ,  $W = 21$ ,  $P < 0.05$ ). Because the perceptual DM, like the cortical DMs, exhibited a clear relationship to stimulus frequency differences, we anticipated finding significant relationships between the perceptual and cortical DMs. Indeed, Spearman correlation and Wilcoxon signed-rank tests (Fig. 6C) revealed that the perceptual space occupied by the auditory stimuli was significantly correlated to the representational space in bilateral AR<sub>eroded</sub> regardless of mask (auditory mask: mean  $r = 0.39$ ,  $W = 55$ ,  $P < 0.001$ ; tactile mask: mean  $r = 0.45$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile mask: mean  $r = 0.41$ ;  $W = 55$ ;  $P < 0.001$ ). More interestingly, the perceptual and cortical DMs were also significantly correlated for all masks in bilateral SR<sub>eroded</sub> (Fig. 6C), although they tended to be stronger correlations with the auditory mask (auditory mask: mean  $r = 0.41$ ,  $W = 55$ ,  $P < 0.001$ ; tactile mask: mean  $r = 0.22$ ;  $W = 46$ ;  $P < 0.05$ ; audio-tactile mask: mean  $r = 0.21$ ;  $W = 45$ ;  $P < 0.05$ ). These results indicate that auditory frequency representations in both parietal and temporal lobe regions relate to auditory perception. Notably, with the auditory mask, the strengths of the correlations between the perceptual DM and the parietal and temporal





**Figure 6.** Representational similarity analysis results for bilateral somatosensory and auditory ROI. (A) *Top*: Dissimilarity matrix (DM) for bilateral AR<sub>eroded</sub> and SR<sub>eroded</sub> for voxels identified in the auditory localizer mask (Materials and Methods). DM entries depict distance metrics that quantify the dissimilarity of ROI spatial response patterns associated with all pairwise combinations of the tested auditory frequencies (F1, F2, F3, F4, and F5 correspond to 75, 130, 195, 270, and 355 Hz, respectively). For visualization purposes, distance metrics are scaled from 0 (most similar) to 1 (most dissimilar). *Bottom*: The perceptual DM indicates the perceived dissimilarity of each frequency pair averaged over subjects (Materials and Methods). (B) Dissimilarity metrics (unscaled) for each somatosensory and auditory ROI are plotted as a function of the absolute difference in frequency between stimuli comprising each pairwise comparison. Regression lines indicate the linear relationship between dissimilarity and frequency difference for group-averaged data. Data and regression lines are plotted separately for the auditory (A; red), tactile (T; green), and audio-tactile (AT; blue) localizer masks. Asterisks indicate significant linear relationships after FDR correction. (C) Average Spearman correlation coefficients across subjects indicating the relationship between the group perceptual DM and the cortical DM for each localizer mask in each ROI. Asterisks denote significant correlations (FDR corrected) and error bars depict bootstrapped 95% confidence intervals. (D) Average Spearman correlation coefficients across subjects indicating the relationship between the perceptual and cortical DMs for each localizer mask in each ROI as a function of ROI size (10–400 voxels in increments of 10 voxels; Materials and Methods). Asterisks denote significant correlations (FDR corrected) and error bars depict bootstrapped 95% confidence intervals.

cortical DMs are nearly matched. While these analyses were performed on ROI voxels exhibiting positive BOLD signal changes in the localizer scans (Materials and Methods), in supplemental analyses we found that patterns over ROI voxels that deactivated in the localizer scans ( $t < 0$ ) did not relate significantly to auditory stimulus frequency.

To test whether these results depended on the number of voxels included in the multivariate analyses (Fig. 6D), we repeated the RSA for bilateral AR<sub>eroded</sub> and SR<sub>eroded</sub> over a range of ROI sizes after first sorting the voxels comprising the unimodal and audio-tactile masks by  $t$ -values computed in the localizer analyses. Voxel selection was performed for each subject separately. Voxel inclusion depended on the magnitude of their  $t$ -values—the smallest ROIs included the voxels with the largest  $t$ -values and the expansion of ROIs resulted from the inclusion of voxels with progressively smaller  $t$ -values. With AR<sub>eroded</sub>, significant correlations between the perceptual and cortical DMs were observed regardless of mask over nearly the entire range of tested ROI sizes (ranging from 10 to 400 voxels). Correlations generally improved with larger ROIs, but the inclusion of additional voxels beyond 150–200 voxels minimally improved correlations. With SR<sub>eroded</sub>, significant correlations were attained with voxels from the auditory localizer mask for ROIs exceeding 130 voxels. In contrast,

correlations between the perceptual and cortical DMs with the tactile and audio-tactile masks showed a stronger dependence on ROI size. With the tactile mask, significant correlations required between 50 and 200 voxels. Beyond this size, the inclusion of more voxels weakened the statistical relationship between the perceptual and cortical DMs. With the audio-tactile mask, significant correlations were only obtained with ROIs including 300–350 voxels. These collective results indicate that auditory frequency representations in bilateral SR<sub>eroded</sub> are sensitive to ROI size and voxel inclusion criteria, unlike those in AR<sub>eroded</sub> which are more robust to the same criteria. Furthermore, the results for SR<sub>eroded</sub> imply that subgroups of voxels in classically defined somatosensory regions of the parietal lobe respond to audition and are better suited to represent auditory frequency information.

#### RSA on Data Separated by Hemisphere

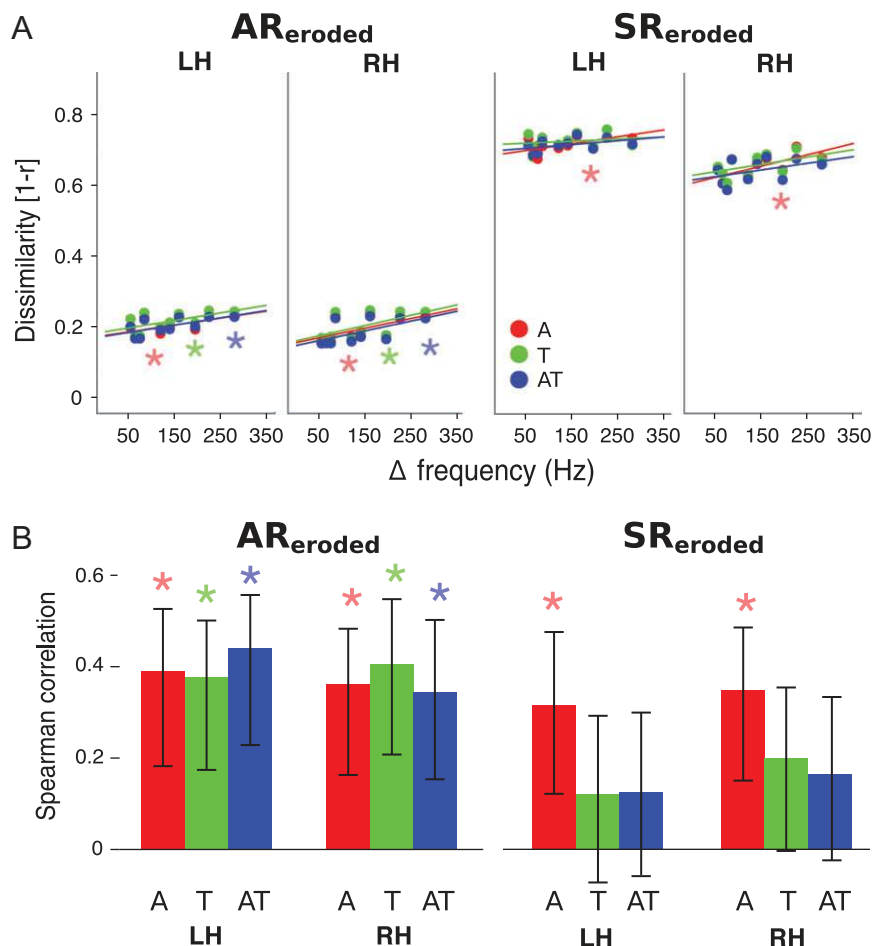
By characterizing auditory frequency representations in the spatial activity patterns over general bilateral somatosensory and auditory ROIs, potential differences between the 2 hemispheres may have been obscured. To explore laterality effects, we performed the RSA for the left and right hemispheres separately (Fig. 7). For AR<sub>eroded</sub>, the positive relationship between

frequency distance and pattern dissimilarity (Fig. 7A) was significant in both the left and right hemispheres irrespective of mask (auditory<sub>LH</sub>: mean  $r = 0.48$ ,  $W = 55$ ,  $P < 0.001$ ; tactile<sub>LH</sub>: mean  $r = 0.47$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile<sub>LH</sub>: mean  $r = 0.51$ ;  $W = 55$ ;  $P < 0.001$ ; auditory<sub>RH</sub>: mean  $r = 0.54$ ,  $W = 55$ ,  $P < 0.001$ ; tactile<sub>RH</sub>: mean  $r = 0.55$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile mask<sub>RH</sub>: mean  $r = 0.52$ ;  $W = 55$ ;  $P < 0.001$ ). Additionally, the AR<sub>eroded</sub> cortical DMs for both hemispheres were significantly correlated with the perceptual DM with all masks (Fig. 7B) (auditory<sub>LH</sub>: mean  $r = 0.4$ ,  $W = 55$ ,  $P < 0.001$ ; tactile<sub>LH</sub>: mean  $r = 0.38$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile<sub>LH</sub>: mean  $r = 0.44$ ;  $W = 55$ ;  $P < 0.001$ ; auditory<sub>RH</sub>: mean  $r = 0.36$ ,  $W = 55$ ,  $P < 0.001$ ; tactile mask<sub>RH</sub>: mean  $r = 0.4$ ;  $W = 55$ ;  $P < 0.001$ ; audio-tactile mask<sub>RH</sub>: mean  $r = 0.34$ ;  $W = 55$ ;  $P < 0.001$ ). Thus, dividing the data in auditory regions by hemisphere did not substantially impact the frequency representation patterns or their relationship with perception for any of the masks. In contrast, separating the data according to hemisphere disrupted the relationship between the multivariate representations in SR<sub>eroded</sub> and the perceptual DM for some of the masks. Although the positive correlation between frequency distance and pattern dissimilarity remained for the somatosensory ROI in the left and right hemispheres with all masks (Fig. 7A), the correlation achieved

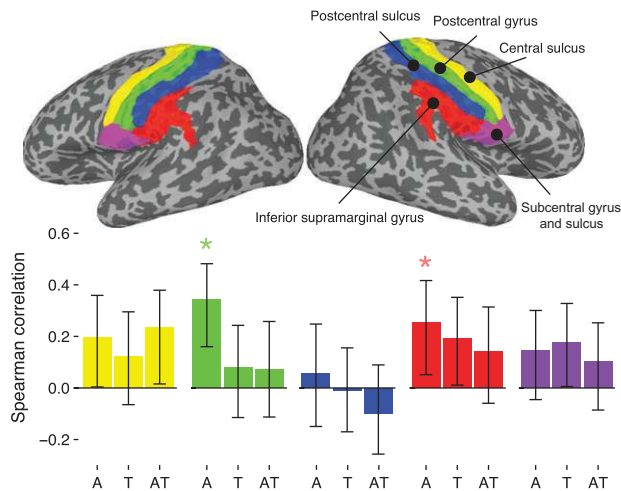
statistical significance only with the auditory masks (auditory<sub>LH</sub>: mean  $r = 0.24$ ,  $W = 45$ ,  $P < 0.05$ ; tactile<sub>LH</sub>: mean  $r = 0.07$ ;  $W = 34$ ;  $P = 0.25$ ; audio-tactile<sub>LH</sub>: mean  $r = 0.1$ ;  $W = 38$ ;  $P = 0.15$ ; auditory<sub>RH</sub>: mean  $r = 0.37$ ,  $W = 54$ ,  $P < 0.002$ ; tactile mask<sub>RH</sub>: mean  $r = 0.18$ ;  $W = 41$ ;  $P = 0.1$ ; audio-tactile mask<sub>RH</sub>: mean  $r = 0.18$ ;  $W = 46$ ;  $P < 0.05$ ). Accordingly, significant correlations between the left and right SR<sub>eroded</sub> cortical DMs and the perceptual DM were observed only with the auditory masks (Fig. 7B) (auditory<sub>LH</sub>: mean  $r = 0.31$ ,  $W = 54$ ,  $P < 0.002$ ; tactile<sub>LH</sub>: mean  $r = 0.12$ ;  $W = 42.5$ ;  $P = 0.07$ ; audio-tactile<sub>LH</sub>: mean  $r = 0.12$ ;  $W = 37$ ;  $P = 0.2$ ; auditory<sub>RH</sub>: mean  $r = 0.35$ ,  $W = 54$ ,  $P < 0.002$ ; tactile mask<sub>RH</sub>: mean  $r = 0.2$ ;  $W = 45$ ;  $P < 0.05$ ; audio-tactile mask<sub>RH</sub>: mean  $r = 0.16$ ;  $W = 42.5$ ;  $P = 0.07$ ).

#### RSA on Parietal Cortex Subdivisions

To determine whether auditory frequency representations were broadly distributed over parietal regions or if they were more spatially organized across the putative somatosensory cortical hierarchy, we performed RSA on data divided according to narrower anatomical ROI definitions (Materials and Methods). While positive relationships between the cortical and perceptual DMs were seen in nearly all of the bilateral parietal ROIs, significant correlations were only observed in the



**Figure 7.** Representational similarity analysis results for unilateral somatosensory and auditory ROI in the left and right hemispheres (LH and RH, respectively) are plotted as a function of the absolute difference in frequency between stimuli comprising each pairwise comparison. Conventions as in Figure 5B. (B) Average Spearman correlation coefficients across subjects indicating the relationship between the perceptual and cortical DMs for each localizer mask in each ROI. Asterisks denote significant correlations (FDR corrected) and error bars depict bootstrapped 95% confidence intervals.



**Figure 8.** Representational similarity analysis results for bilateral parietal regions sorted according to anatomical subdivisions. Spearman correlation coefficients across subjects indicating the relationship between the perceptual and cortical DMs for each localizer mask in each ROI. Separate bilateral ROIs were created by intersecting each localizer mask (auditory, tactile, and audio-tactile) with anatomical subdivision masks (central sulcus, postcentral gyrus, postcentral sulcus, supramarginal gyrus, subcentral gyrus/sulcus; Materials and Methods). Asterisks denote significant correlations (FDR corrected) and error bars depict bootstrapped 95% confidence intervals.

postcentral gyrus and the supramarginal gyrus (Fig. 8) and these significant patterns were only evident with the auditory mask. These results demonstrate that auditory representations in parietal cortex are not restricted to the higher-order somatosensory regions in close proximity to auditory cortex, as would be expected if the parietal responses were simply due to partial volume effects. Additionally, they confirm that the parietal lobe voxels that responded best to sounds in the localizer scans were the voxels most likely to represent auditory frequency information in a manner consistent with perception.

## Discussion

Previous studies have characterized human cortical responses to auditory and tactile stimulation and reported extensive overlap in temporal lobe regions traditionally considered to support auditory functions. These studies found that responses in primary and association auditory cortex could be modulated or driven by tactile stimulation. We elaborated on previous work in multiple ways. First, we focused on responses to auditory stimulation and tested for the possibility that activity in brain regions outside of the temporal lobes reflects auditory signal processing. Second, we systematically manipulated the frequency content of the auditory stimuli to characterize frequency-dependent responses in parietal regions. Third, we employed univariate and multivariate analysis approaches to characterize the relationship between brain activity patterns and auditory frequency. Finally, we related the representational space occupied by auditory stimuli in the BOLD fMRI responses to the representational space defined by participants' perception of the same stimuli. Our primary result is that BOLD activation patterns in bilateral parietal lobe regions, in addition to expected temporal lobe areas, related to auditory frequency information in a manner that corresponded to offline perceptual similarity judgments. This novel finding suggests that auditory frequency information is not confined to classically defined auditory cortex, but is also distributed across classically defined somatosensory areas.

We found that auditory stimulation modulated BOLD signals in the parietal lobe with stereotypical hemodynamic response profiles (Fig. 2). A previous report focusing on enhanced auditory responses in the parietal operculum in an individual with sound-touch synesthesia also found auditory responses, of comparable magnitude to the responses we observed, in the parietal operculum of normal control participants (Beauchamp and Ro 2008). The results of our surface-based univariate analysis are consistent with these earlier findings, showing an activation cluster in the parietal operculum with a peak in the parietal operculum subdivision OP4 (Eickhoff et al. 2006). The parietal operculum can also be selectively active in patients experiencing tinnitus (Job et al. 2012), the perception of auditory tones in the absence of external stimuli, and in control subjects experiencing transient tinnitus-like effects (Job et al. 2014). Furthermore, a recent fMRI decoding study reported that multivariate activation patterns in putative primary somatosensory cortex could be uniquely associated with stimulation in a number of sensory modalities including audition (Liang et al. 2013). Our results extend on these findings by showing that the magnitude of BOLD signal responses in the parietal lobe varies with frequency and that multivariate activation patterns systematically relate to auditory perception. Moreover, by measuring responses to auditory, tactile, and audio-tactile stimulation separately in the localizer scans, we found that these sensory conditions evoked responses in distinct but partially overlapping voxel groups. These differences presumably reflect the localization of intermixed neural populations that are preferentially responsive to tactile or auditory stimulation. Consistent with this, both univariate and multivariate analyses of the discrimination scan data showed that the parietal cortex voxels comprising the auditory localizer mask contained more robust auditory frequency representations as compared to voxels comprising the tactile and audio-tactile masks. These results suggest that auditory frequency representations in parietal regions may be carried in neural subpopulations that are distributed in a distinct pattern compared to those that are responsive to somatosensory inputs. These observations need to be confirmed in ultra-high resolution scanning and single-unit recording experiments testing a more comprehensive range of tones.

Although our analyses revealed auditory responses distributed over both dorsal and ventral portions of the parietal lobe, we were particularly careful in defining analysis masks in perisylvian cortex. The Sylvian fissure is straddled by classically defined secondary somatosensory cortex (in the parietal operculum) and auditory cortex (in the superior temporal gyrus). This arrangement makes localizing perisylvian responses challenging because of partial volume effects (Ozcan et al. 2005). We addressed this concern in multiple ways. First, we acquired our data in 2-mm isotropic voxels and we avoided spatial smoothing while maintaining the data in native volume space whenever possible. Second, when we spatially filtered or sampled our data, we performed these operations in surface space rather than volume space to reduce cross-fissure contamination. Third, we carefully inspected the spatial response profiles in the parietal operculum to confirm the presence of auditory responses far from the areal boundaries (Fig. 2). Fourth, we adopted conservative ROI border definitions that minimized the inclusion of ambiguous voxels. Indeed, ROIs eroded by 2 mm to account for potential partial volume effects at the borders between the parietal and temporal lobes (Materials and Methods) contained clear auditory responses and frequency representations (Figs. 5 and 6). In control analyses, we repeated the RSA using more extreme erosion levels (Supplementary Fig. 1) and found nearly identical

results with ROIs eroded by 4 mm. RSA with the auditory mask revealed significant auditory representations in somatosensory regions even when performed on a ROI eroded by 10 mm (Supplementary Fig. 1). These supplementary analyses confirm that auditory information can be represented in parietal cortex even when the perisylvian regions most susceptible to partial volume effects or spatial blurring artifacts, particularly those in the parietal operculum, are excluded.

Stimulus control is especially important when interpreting crossmodal activations and we took multiple steps to ensure that our results could not be trivially explained by confounds in our stimulation paradigms. First, we confirmed that the tactile stimuli used during the fMRI scans were inaudible to the subjects in control experiments. This confirmation is important because the distinction between the auditory and tactile analysis masks would be confounded if the vibrations occurring on the fingers were explicitly associated with sounds. In control experiments performed in the scanner, participants achieved only chance-level performance on a tactile detection task in the absence of physical contact with the tactile stimulator (mean detection rate:  $0.49 \pm 0.012$ )—presumably, performance would have been better had the subjects been able to hear the vibrations. Second, we addressed the concern that the parietal responses to auditory stimulation reflected tactile processing explicitly. Specifically, because the MRI-compatible in-ear headphones used in the fMRI scans produce sounds via vibrating piezoelectric elements, cutaneous receptors in the ear canal may have been mechanically driven by the earbuds to produce the responses in somatosensory regions. A number of considerations argue against this concern. First, because of the somatotopic organization of the somatosensory system in parietal cortex, the sound-induced activation patterns would presumably be spatially constrained to only the ear representations if they were associated with cutaneous receptor activation in the ear canal. Our data indicate that auditory responses are broadly distributed over the parietal lobe. Second, if the sound-related responses were due to cutaneous activation, the bimodal response profile over auditory frequencies in  $SR_{eroded}$  (Fig. 5B) would presumably be reflected in frequency response profiles associated with actual tactile stimulation. Notably, we find that tactile stimulation at vibration frequencies matching the sound frequencies tested in the auditory discrimination scans produced response profiles that clearly differed from the pattern observed with sounds (data not shown). Lastly, in control fMRI experiments designed specifically to test whether vibrating earbud elements are required for sound-related parietal activations we found that significant BOLD signal changes in parietal brain regions can be achieved using sound stimuli delivered via circumaural, air-conduction headphones (Supplementary Fig. 2). Collectively, our main and Supplementary data suggest that the activation patterns we observe in parietal cortex relate to auditory processing. Importantly, the control fMRI experiment also served to demonstrate that auditory responses could be observed in parietal regions even in scan sessions that involved no tactile stimulation. This addresses concerns that the auditory activations in somatosensory cortex during the main experiments resulted from a rapidly learned association between the sounds and vibrations experienced during the localizer and discrimination scans. Together, these supplemental results imply that the activation patterns to sound stimulation in parietal cortex cannot be trivially attributed to any particular stimulation paradigms and experimental contexts.

RSA revealed that auditory stimulus frequency is associated with distinct multi-voxel activation patterns distributed over bilateral portions of parietal cortex. These analyses revealed a

relationship between the distributed activation patterns and auditory frequency perception. Specifically, the representational space occupied by the different auditory stimuli, defined by the similarity of their multi-voxel response patterns, was significantly correlated with the perceived similarity of the tones. In other words, stimuli that were perceived as more similar tended to associate with more similar spatial activation patterns. As expected, this relationship was true of activation patterns in bilateral superior temporal lobe regions, which effectively served as positive control ROIs. Surprisingly, the strengths of the correlations between the perceptual and cortical DMs were similar for the auditory and somatosensory ROI using some localizer masks. In (negative) control analyses, we performed RSA on multivariate patterns measured in a visual cortex region (Supplementary Fig. 3) and failed to find significant correlations between the cortical representational spaces and the perceptual DM. Thus, the relationship between cortical and perceptual DMs for the auditory stimuli tested in our experiments appears to be specific to temporal and parietal cortex. Admittedly, our approach can be substantially improved in future studies. For instance, we compared the cortical dissimilarity measures to the perceptual dissimilarity judgments tested offline and averaged over subjects. We took this approach because the task performed during the discrimination scans, which was designed to ensure that participants attended to stimulus frequency, did not permit a sensitive analysis of auditory discrimination performance. Future experiments to confirm the relationship between parietal lobe responses and auditory perception should be designed to assess subject-specific and temporally-specific (i.e., trial-wise) associations.

How might auditory information reach the parietal cortex? Extensive cortico-cortical and thalamo-cortical pathways connecting the somatosensory and auditory systems have been identified using invasive tracing methods (Cappe and Barone 2005; Hackett et al. 2007; Cappe et al. 2009) and non-invasive structural imaging in humans (Ro et al. 2013). This anatomical connectivity could account for the intrinsic connectivity between somatosensory and auditory cortex calculated from correlated spontaneous BOLD signal fluctuations (Power et al. 2011). Notably, specific subdivisions of the auditory cortical system may be more related to parietal cortex processing than others. Indeed, cortical regions comprising the putative auditory dorsal pathway (Rauschecker and Tian 2000) are responsive to touch (Fuxe et al. 2002; Fu et al. 2003; Schurmann et al. 2006) and exhibit strong connectivity with parietal and frontal regions (Romanski et al. 1999; Kaas and Hackett 2000). Importantly, these caudal belt areas, which are thought to support spatial perception and audio-motor processing (Rauschecker and Scott 2009; Rauschecker 2011), contain neurons that exhibit clear frequency selectivity (Kusmierek and Rauschecker 2014) and could thus contribute to the auditory frequency responses we observed as well as the responses observed in portions of posterior parietal and prefrontal cortex that appear to be specialized for auditory spatial processing (Bushara et al. 1999). Due to the limited temporal resolution of fMRI, it is unclear whether the auditory responses in parietal cortex reflect feedforward or feedback signaling. Auditory inputs could conceivably reach somatosensory regions with relatively short latencies and laminar response profiles consistent with feedforward processing as is the case with somatosensory inputs to some auditory regions (Schroeder and Foxe 2002). Alternatively, but not exclusively, the auditory responses in parietal regions could reflect feedback projections from frontal regions that support working memory and decision-making independent of sensory modality (Vergara et al. 2016).

Neural associations between audition and touch likely relate to the ways in which the 2 sensory modalities overlap in

information processing. Speech perception and production clearly involve both senses (Ito et al. 2009). Our ability to judge surface texture also recruits both audition and touch in order to process the sounds and mechanical vibrations produced when we palpate objects, respectively (Lederman 1979; Yau et al. 2009a). Notably, signals in the 100–300 Hz ranges are critical to both speech (e.g., the fundamental frequencies of human voices; Lattner et al. 2005) and texture processing (Manfredi et al. 2014), thus it is unsurprising that audition and touch exhibit online and offline perceptual interactions at these frequencies (Yau et al. 2009b; Crommett et al. 2017; Yau et al. 2010). Interestingly, although common neural circuits may support frequency processing for audition and touch (Butler et al. 2012), processing of other temporal features, like stimulus duration (Butler et al. 2011), appears to rely on modality-dedicated circuits which implies that the neural associations between audition and touch may be feature-specific. The co-occurrence of auditory and somatosensory signals that are highly specific in the frequency domain may build audio-tactile associations in frequency-processing circuits through an extensive history of co-activation. In this framework, the auditory responses in parietal cortex could reflect predictive coding mechanisms in which hearing specific tones evokes neural representations of the tactile events associated with the sounds even in the absence of mechanical stimulation. Accordingly, parietal regions would only be responsive to a relatively low range of auditory frequencies which overlap those experienced through touch. This key prediction should be tested in future experiments.

Previous efforts to characterize the neural associations between audition and touch have primarily focused on the capacity for somatosensory stimulation to evoke or modulate activity in auditory cortex. The present data reveal that brain regions traditionally included in the somatosensory cortical hierarchy can also respond to auditory stimulation. Across multiple parietal cortex subdivisions, we found that BOLD response magnitude depended on auditory stimulus frequency. Moreover, we observed that auditory stimuli associated with distinct multi-voxel activation patterns and the similarity of these response patterns related to how similarly the tones were perceived. Our data demonstrate that parietal cortex can signal auditory frequency, but a number of important questions regarding frequency representations in the parietal lobe remain to be addressed. How do neurons residing in parietal regions represent frequency signals? In non-human primates, neurons in primary somatosensory cortex can signal vibration frequencies above 100 Hz in their spike timing (Harvey et al. 2013), but evidence for explicit frequency tuning in somatosensory cortex remains elusive. Whether auditory information is similarly carried in somatosensory cortex using a temporal code is unknown. Additionally, it is unclear whether frequency representations are topographically organized in the parietal cortex the way they are in multiple fields in auditory cortex. Despite these open questions, our data provide preliminary evidence that auditory frequency information can be distributed over a vast network of brain regions spanning parietal as well as temporal cortex. Future research needs to address the functional roles of these regions and how they collaborate to support perception and cognition.

## Supplementary Material

Supplementary data are available at *Cerebral Cortex* online.

## Funding

This work was supported by the National Institutes of Health (R01NS097462 to J.M.Y.), Alfred P. Sloan Research Fellowship (to J.M.Y.), and the Caroline Wiess Law Fund for Research in Molecular Medicine (to J.M.Y.).

## Notes

We would like to thank R. Savjani, D. Ress, M. Beauchamp, and N. Oosterhof for thoughtful discussions, A. Lin and K. Runge for help with data collection, M. Tommerdahl and J. Holden for technical assistance with the CM3. We acknowledge the Core for Advanced MRI (CAMRI). *Conflict of Interest*: None declared.

## References

- Beauchamp MS, Ro T. 2008. Neural substrates of sound-touch synesthesia after a thalamic lesion. *J Neurosci.* 28:13696–13702.
- Butler JS, Foxe JJ, Fiebelkorn IC, Mercier MR, Molholm S. 2012. Multisensory representation of frequency across audition and touch: high density electrical mapping reveals early sensory-perceptual coupling. *J Neurosci.* 32:15338–15344.
- Butler JS, Molholm S, Fiebelkorn IC, Mercier MR, Schwartz TH, Foxe JJ. 2011. Common or redundant neural circuits for duration processing across audition and touch. *J Neurosci.* 31:3400–3406.
- Bushara KO, Weeks RA, Ishii K, Catalan M, Tian B, Rauschecker JP, Hallett M. 1999. Modality-specific frontal and parietal areas for auditory and visual spatial localizations in humans. *Nat Neurosci.* 2:759–766.
- Cappe C, Barone P. 2005. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur J Neurosci.* 22:2886–2902.
- Cappe C, Rouiller EM, Barone P. 2009. Multisensory anatomical pathways. *Hear Res.* 258:28–36.
- Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 29:162–173.
- Crommett LE, Perez-Bellido A, Yau JM. 2017. Auditory adaptation improves tactile frequency perception. *J Neurophysiol.* 117:1352–1362.
- Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage.* 53:1–15.
- Driver J, Noesselt T. 2008. Multisensory interplay reveals cross-modal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron.* 57:11–23.
- Eickhoff SB, Schleicher A, Zilles K, Amunts K. 2006. The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. *Cereb Cortex.* 16:254–267.
- Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex.* 1:1–47.
- Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage.* 9:195–207.
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. 2002. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol.* 88:540–543.
- Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE. 2003. Auditory cortical neurons

- respond to somatosensory stimulation. *J Neurosci.* 23: 7510–7515.
- Ghazanfar AA, Schroeder CE. 2006. Is neocortex essentially multisensory? *Trends Cogn Sci.* 10:278–285.
- Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, de la Mothe LA, Schroeder CE. 2007. Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception.* 36:1419–1430.
- Harvey MA, Saal HP, Dammann JF 3rd, Bensmaia SJ. 2013. Multiplexing stimulus information through rate and temporal codes in primate somatosensory cortex. *PLoS Biol.* 11:e1001558.
- Ito T, Tiede M, Ostry DJ. 2009. Somatosensory function in speech perception. *Proc Natl Acad Sci USA.* 106:1245–1248.
- Job A, Jacob R, Pons Y, Raynal M, Kossowski M, Gauthier J, Lombard B, Delon-Martin C. 2014. Specific activation of opercular 3 (OP3) brain region during provoked tinnitus-related phantom auditory perceptions in humans. *Brain Struct Funct.* 221:912–922.
- Job A, Pons Y, Lamalle L, Jaillard A, Buck K, Segebarth C, Delon-Martin C. 2012. Abnormal cortical sensorimotor activity during “Target” sound detection in subjects with acute acoustic trauma sequelae: an fMRI study. *Brain Behav.* 2:187–199.
- Johnson KO. 2001. The roles and functions of cutaneous mechanoreceptors. *Curr Opin Neurobiol.* 11:455–461.
- Jousmaki V, Hari R. 1998. Parchment-skin illusion: sound-biased touch. *Curr Biol.* 8:R190.
- Kaas JH, Hackett TA. 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci USA.* 97: 11793–11799.
- Kayser C, Petkov CI, Augath M, Logothetis NK. 2005. Integration of touch and sound in auditory cortex. *Neuron.* 48:373–384.
- Kusmierek P, Rauschecker JP. 2014. Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *J Neurophysiol.* 111:1671–1685.
- Lattner S, Meyer ME, Friederici AD. 2005. Voice perception: sex, pitch, and the right hemisphere. *Hum Brain Mapp.* 24:11–20.
- Lederman SJ. 1979. Auditory texture perception. *Perception.* 8: 93–103.
- Lemus L, Hernandez A, Luna R, Zainos A, Romo R. 2010. Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron.* 67:335–348.
- Liang M, Mouraux A, Hu L, Iannetti GD. 2013. Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nat Commun.* 4:1979.
- Manfredi LR, Saal HP, Brown KJ, Zielinski MC, Dammann JF 3rd, Polashock VS, Bensmaia SJ. 2014. Natural scenes in tactile texture. *J Neurophysiol.* 111:1792–1802.
- Moerel M, De Martino F, Formisano E. 2014. An anatomical and functional topography of human auditory cortical areas. *Front Neurosci.* 8:225.
- Mountcastle VB. 2005. *The sensory hand. Neural mechanisms in somatic sensation.* Cambridge (MA): Harvard University Press.
- Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N. 2014. A toolbox for representational similarity analysis. *PLoS Comput Biol.* 10:e1003553.
- Nordmark PF, Pruszynski JA, Johansson RS. 2012. BOLD responses to tactile stimuli in visual and auditory cortex depend on the frequency content of stimulation. *J Cogn Neurosci.* 24:2120–2134.
- Occelli V, Spence C, Zampini M. 2011. Audio-tactile interactions in temporal perception. *Psychon Bull Rev.* 18:429–454.
- Ozcan M, Baumgartner U, Vucurevic G, Stoeter P, Treede RD. 2005. Spatial resolution of fMRI in the human parasyllian cortex: comparison of somatosensory and auditory activation. *Neuroimage.* 25:877–887.
- Pannunzi M, Perez-Bellido A, Pereda-Banos A, Lopez-Moliner J, Deco G, Soto-Faraco S. 2015. Deconstructing multisensory enhancement in detection. *J Neurophysiol.* 113:1800–1818.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. 2011. Functional network organization of the human brain. *Neuron.* 72:665–678.
- Rauschecker JP. 2011. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear Res.* 271:16–25.
- Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci.* 12:718–724.
- Rauschecker JP, Tian B. 2000. Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci USA.* 97:11800–11806.
- Ro T, Ellmore TM, Beauchamp MS. 2013. A neural link between feeling and hearing. *Cereb Cortex.* 23:1724–1730.
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex [see comments]. *Nat Neurosci.* 2:1131–1136.
- Romo R, Salinas E. 2003. Flutter discrimination: neural codes, perception, memory and decision making. *Nat Rev Neurosci.* 4:203–218.
- Rosen BR, Buckner RL, Dale AM. 1998. Event-related functional MRI: past, present, and future. *Proc Natl Acad Sci USA.* 95: 773–780.
- Saad ZS, Reynolds RC. 2012. SUMA. *Neuroimage.* 62:768–773.
- Schroeder CE, Foxe JJ. 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res.* 14:187–198.
- Schurmann M, Caetano G, Hlushchuk Y, Jousmaki V, Hari R. 2006. Touch activates human auditory cortex. *Neuroimage.* 30:1325–1331.
- Vergara J, Rivera N, Rossi-Pool R, Romo R. 2016. A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron.* 89:54–62.
- Wilson EC, Reed CM, Braid LD. 2009. Integration of auditory and vibrotactile stimuli: effects of phase and stimulus-onset asynchrony. *J Acoust Soc Am.* 126:1960–1974.
- Wilson EC, Reed CM, Braid LD. 2010. Integration of auditory and vibrotactile stimuli: effects of frequency. *J Acoust Soc Am.* 127:3044–3059.
- Yau JM, Hollins M, Bensmaia SJ. 2009a. Textural timbre: the perception of surface microtexture depends in part on multimodal spectral cues. *Commun Integr Biol.* 2:344–346.
- Yau JM, Olenczak JB, Dammann JF, Bensmaia SJ. 2009b. Temporal frequency channels are linked across audition and touch. *Curr Biol.* 19:561–566.
- Yau JM, Weber AI, Bensmaia SJ. 2010. Separate mechanisms for audio-tactile pitch and loudness interactions. *Front Psychol.* 1:160. doi:10.3389/fpsyg.2010.00160.