

# Tuning In to Sound: Frequency-Selective Attentional Filter in Human Primary Auditory Cortex

Sandra Da Costa,<sup>1,2</sup> Wietske van der Zwaag,<sup>3</sup> Lee M. Miller,<sup>4,5</sup> Stephanie Clarke,<sup>1</sup> and Melissa Saenz<sup>2,6</sup>

<sup>1</sup>Neuropsychology and Neurorehabilitation Service, Department of Clinical Neurosciences, Lausanne University Hospital, Lausanne, Switzerland 1011, <sup>2</sup>Laboratoire de Recherche en Neuroimagerie, Department of Clinical Neurosciences, Lausanne University Hospital, Lausanne, Switzerland 1011, <sup>3</sup>Center for Biomedical Imaging, University of Lausanne, Lausanne, Switzerland 1015, <sup>4</sup>Center for Mind and Brain, University of California, Davis, Davis, California 95618, <sup>5</sup>Department of Neurobiology, Physiology, and Behavior, University of California, Davis, Davis, California 95616, and <sup>6</sup>Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland 1015

Cocktail parties, busy streets, and other noisy environments pose a difficult challenge to the auditory system: how to focus attention on selected sounds while ignoring others? Neurons of primary auditory cortex, many of which are sharply tuned to sound frequency, could help solve this problem by filtering selected sound information based on frequency-content. To investigate whether this occurs, we used high-resolution fMRI at 7 tesla to map the fine-scale frequency-tuning (1.5 mm isotropic resolution) of primary auditory areas A1 and R in six human participants. Then, in a selective attention experiment, participants heard low (250 Hz)- and high (4000 Hz)-frequency streams of tones presented at the same time (dual-stream) and were instructed to focus attention onto one stream versus the other, switching back and forth every 30 s. Attention to low-frequency tones enhanced neural responses within low-frequency-tuned voxels relative to high, and when attention switched the pattern quickly reversed. Thus, like a radio, human primary auditory cortex is able to tune into attended frequency channels and can switch channels on demand.

## Introduction

The “cocktail party” problem (Cherry, 1953) refers to the challenge of auditory selective attention: how to focus attention onto selected sounds in a noisy background? In studies of visual attention, much evidence points toward a “feature-based” mechanism by which attention to a particular visual feature enhances the response of visual cortical neurons tuned to that feature, thus strengthening the neural representation of attended stimuli relative to unattended stimuli (Treue and Martinez-Trujillo, 1999; Saenz et al., 2002; Maunsell and Treue, 2006). In the auditory system, the most ubiquitous feature to which cortical neurons are tuned is sound frequency. Here we test whether frequency-tuned units of human primary areas A1 and R are modulated by selective attention to preferred versus nonpreferred sound frequencies in the dynamic manner needed to account for human listening abilities. Such an early-stage filtering mechanism could contribute to downstream selection of spectrally complex auditory stimuli like speech.

Previous human studies suggest that attention modulates responses in the region of primary auditory cortex (EEG: Hillyard

et al., 1973; Woods et al., 1984; Woldorff et al., 1993; MEG: Fujiwara et al., 1998 fMRI: Jäncke et al., 1999; Rinne et al., 2008; EcoG: Bidet-Caulet et al., 2007) including frequency-specific enhancement (Paltoglou et al., 2009; Oh et al., 2012). Other fMRI studies suggest that attentional modulation occurs predominantly in secondary, and not primary, auditory cortical areas (Petkov et al., 2004; Woods et al., 2009; Woods et al., 2010; Ahveninen et al., 2011). Differences across studies may relate to the variety of spatial, featural, and multisensory attentional tasks used. However, previous human studies have not performed fine-scale frequency mappings needed to identify A1 and R in individual subjects (as we aim to do here with high-resolution fMRI).

Single-neuron recordings in A1 of rats and ferrets show that attention to a target tone amid distractor sounds reshapes the frequency-tuning profiles of individual neurons (Fritz et al., 2003; Fritz et al., 2005; Atiani et al., 2009; Jaramillo and Zador, 2011; David et al., 2012). While the specific modulatory effects vary (see Discussion), attention tends to enhance the contrast between responses to target and non-target frequencies. One caveat is that the animals require many weeks of specific task training and thus the effects potentially involve long-term learning mechanisms, in addition to the flexible and transient attentional mechanisms needed to account for dynamic human listening skills.

Here, we test for attentional modulation of frequency-tuned units in human primary auditory cortex using a two-step approach. First, we use high-resolution fMRI at 7T to map the fine-scale frequency tuning of human primary auditory areas hA1 and hR in individual subjects. Second, we test whether

Received Sept. 14, 2012; revised Dec. 4, 2012; accepted Dec. 7, 2012.

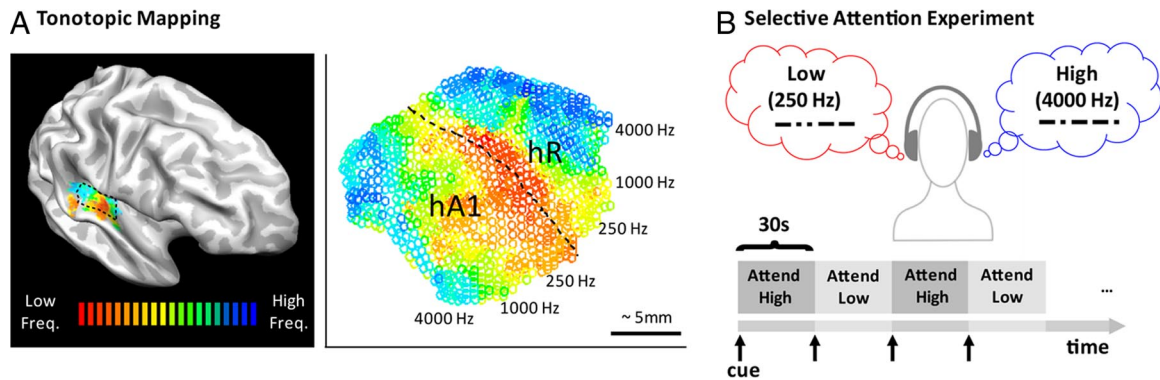
Author contributions: S.D.C., L.M.M., S.C., and M.S. designed research; S.D.C., W.v.d.Z., and M.S. performed research; S.D.C. and M.S. analyzed data; S.D.C., W.v.d.Z., L.M.M., S.C., and M.S. wrote the paper.

This work was supported by Swiss National Science Foundation Grants 320030-124897 to S.C. and IZKOZ3\_139473/1 to Micah Murray, and by the Centre d'Imagerie BioMédicale of the Université de Lausanne, Université de Genève, Hôpitaux Universitaires de Genève, Lausanne University Hospital, Ecole Polytechnique Fédérale de Lausanne, and the Leenaards and Louis-Jeantet Foundations. We thank Micah Murray for hosting the visit of L.M.M., leading to our collaboration.

Correspondence should be addressed to Melissa Saenz, Department of Clinical Neurosciences, Lausanne University Hospital (CHUV), 1011 Lausanne, Switzerland. E-mail: melissa.saenz@epfl.ch.

DOI:10.1523/JNEUROSCI.4405-12.2013

Copyright © 2013 the authors 0270-6474/13/331858-06\$15.00/0



**Figure 1.** *A*, Tonotopic mapping was used to identify primary auditory cortex in each subject individually ( $n = 6$ ). In each hemisphere ( $n = 12$ ), two mirror-symmetric gradients (high-to-low and low-to-high) corresponding to the primary areas hA1 and hR were manually outlined on the medial two-thirds of Heschl's gyrus (one same right hemisphere shown). Each voxel within the selected region was labeled according to its preferred frequency between 88 and 8000 Hz in half-octave steps. *B*, Next, in the selective attention (dual-stream) experiment, low (250 Hz)- and high (4000 Hz)-frequency patterned tonal streams were presented concurrently to different ears. Subjects were cued to attend to only one stream at a time, alternating the attended stream every 30 s (blocks of attend high vs attend low). A 2-IFC experiment was used to focus attention on the cue stream (see Materials and Methods). The stimulus itself did not change across blocks, only the attentional state. Ear-side was counterbalanced across runs allowing the comparison of effects of frequency-specific attention (attend high vs attend low collapsed across sides) to effects of spatial-selective attention (attend contralateral vs attend ipsilateral collapsed across frequencies).

frequency-tuned units are modulated by attention to preferred versus nonpreferred frequencies in a dynamic selective attention task. The results demonstrate robust frequency-specific attentional modulation in primary auditory cortex - these effects outweighed more modest effects of spatial attention and were large relative to stimulus-driven changes.

## Materials and Methods

Six subjects (ages 25–40, 2 males) with no known hearing deficit participated after giving written, informed consent. Experimental procedures were approved by the Ethics Committee of the Faculty of Biology and Medicine of the University of Lausanne.

**MRI data acquisition and data analysis.** Blood oxygenation level-dependent (BOLD) functional imaging was performed with an actively shielded 7 Tesla Siemens MAGNETOM scanner (Siemens Medical Solutions) located at the Centre d'Imagerie BioMedicale in Lausanne, Switzerland. The increased signal-to-noise ratio and available BOLD associated with ultra-high magnetic field systems ( $>3$  T) allow the use of smaller voxel sizes in fMRI. The spatial specificity of the BOLD signal is improved because the signal strength of venous blood is reduced due to a shortened relaxation time, restricting activation signals to cortical gray matter (van der Zwaag et al., 2009, 2011). fMRI data were acquired using an 8-channel head volume RF-coil (RAPID Biomedical GmbH) and a continuous EPI pulse sequence with sinusoidal read-out ( $1.5 \times 1.5$  mm in-plane resolution, slice thickness = 1.5 mm, TR = 2000 ms, TE = 25 ms, flip angle =  $47^\circ$ , slice gap = 1.57 mm, matrix size =  $148 \times 148$ , field of view  $222 \times 222$ , 30 oblique slices covering the superior temporal plane). A T1-weighted high-resolution 3D anatomical image (resolution =  $1 \times 1 \times 1$  mm, TR = 5500 ms, TE = 2.84 ms, slice gap = 1 mm, matrix size =  $256 \times 240$ , field of view =  $256 \times 240$ ) was acquired for each subject using the MP2RAGE pulse sequence optimized for 7T MRI (Marques et al., 2010).

Standard fMRI data preprocessing steps were performed with BrainVoyager QX v2.3 software and included linear trend removal, temporal high-pass filtering, and motion correction. Spatial smoothing was not applied. Functional time-series data were interpolated into  $1 \times 1 \times 1$  mm volumetric space and registered to each subject's 3D Talairach-normalized anatomical dataset. Cortical surface meshes were generated from each subject's anatomical dataset using automated segmentation tools in BrainVoyager QX.

**Auditory stimuli.** Sound stimuli were generated using MATLAB and the Psychophysics Toolbox (www.psychtoolbox.org) with a sampling rate of 44.1 kHz. Stimuli were delivered via MRI-compatible headphones (AudioSystem, NordicNeuroLab) featuring flat frequency transmission over the stimulus range. Sound intensities were adjusted to match stan-

dard equal-loudness curves (ISO 226) at phon 85: the sound intensity of each pure tone stimulus (ranging from 88 to 8000 Hz) was adjusted to approximately equal the perceived loudness of a 1000 Hz reference tone at 85 dB SPL (range of sound intensities: 82–97 dB SPL). Sound levels were further attenuated ( $\sim 22$  dB) by protective ear plugs. Subjects reported hearing sounds clearly over background scanner noise and were instructed to keep eyes closed during fMRI scanning.

**Tonotopic mapping step.** For tonotopic mapping, pure tones (88, 125, 177, 250, 354, 500, 707, 1000, 1414, 2000, 2828, 4000, 5657, and 8000 Hz; half-octave steps) were presented in ordered progressions, following our previously described methods (Da Costa et al., 2011). Briefly, starting with the lowest (or highest) frequency, pure tone bursts of that frequency were presented for a 2 s block before stepping to the next consecutive frequency until all 14 frequencies had been presented. The 28 s progression was followed by a 4 s silent pause, and this 32 s cycle was repeated 15 times per 8 min scan run. Each subject participated in two scan runs (one low-to-high progression and one high-to-low progression), and resulting maps of the two runs were averaged. The frequency progressions were designed to induce a traveling wave of response across cortical tonotopic maps (Engel, 2012). Linear cross-correlation was used to determine the time-to-peak of the fMRI response wave on a per-voxel basis, and to thus assign a corresponding best frequency value to each voxel. Analyses were performed in individual-subject volumetric space and results were then projected onto same-subject cortical surface meshes.

As shown in Figure 1*A*, two tonotopic gradients with mirror symmetry ("high-low-low-high") were clearly observed running approximately across Heschl's gyrus in both hemispheres of all subjects (Formisano et al., 2003; Woods et al., 2009; Humphries et al., 2010; Da Costa et al., 2011; Striem-Amit et al., 2011; Langers and van Dijk, 2012), the more posterior "high-to-low" gradient corresponding to human A1 (hA1) and the more anterior "low-to-high" gradient corresponding to hR. In macaque auditory cortex, fields A1 and R receive parallel thalamic input and are both considered part of the primary auditory core (along with a possible third, smaller field, RT, which has not yet been reliably confirmed in the human), and the relative functions of the two fields remain unknown (Hackett, 2011). We manually outlined a contiguous patch of cortical surface containing the two primary gradients corresponding to hA1 and hR using drawing tools within BrainVoyager QX, as illustrated with dotted lines (Fig. 1*A*). The exact borders were not dependent upon the particular correlation threshold used for display since the overall pattern was observable across a large range of display thresholds. Anterior and posterior borders were drawn along the outer high-frequency representations. Lateral and medial borders were conservatively drawn to include only the medial two-thirds of Heschl's gyrus, in accordance with human architectonics (Rivier and Clarke, 1997; Hackett, 2011). Tonotopic re-

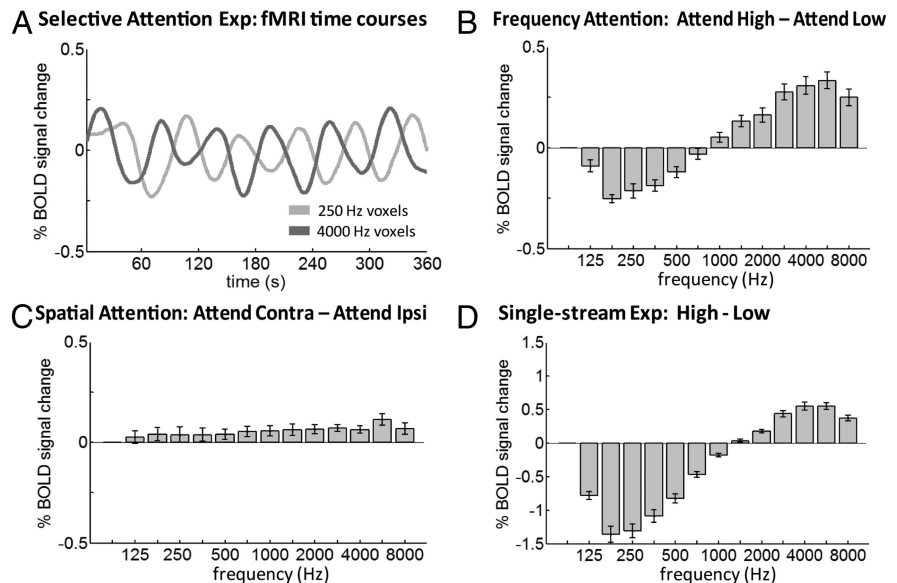
sponses extending onto the lateral end of Heschl's gyrus may include non-primary belt regions. The border between hA1 and hR was drawn across the length of the low-frequency gradient reversal.

Once selected on the cortical surface, the hA1 and hR regions were projected into the same-subject's  $1 \times 1 \times 1$  mm interpolated volumetric space to generate 3D regions of interest (ROIs). The 3D ROIs were generated with a width of 3 mm ( $-1$  mm to  $+2$  mm from the white/gray matter boundary). All volumetric voxels ( $1 \times 1 \times 1$  mm interpolated) falling within the 3D ROIs were labeled with a best-frequency map value, and were subsequently analyzed in the selective attention experiment. Data analysis for the selective attention experiment was thus performed in volumetric space without loss of the acquired spatial resolution.

**Selective attention (dual-stream experiment).** During the selective attention experiment (Fig. 1B), the same subjects attended to one of two competing tonal streams presented simultaneously to different ears—one stream consisted of low-frequency tone bursts (250 Hz) and the other, high (4000 Hz). Ear-side, i.e., whether low-frequency tones were presented to the left or right side, was counterbalanced across runs. By design, this allowed the comparison of any effects of frequency-specific attention (“attend high” vs “attend low” collapsed across sides) to effects of spatial-selective attention (“attend contralateral” vs “attend ipsilateral” collapsed across frequencies). Every 30 s, subjects were cued to switch attention from one stream to the other. The brief auditory cue, appearing at the beginning of each block, was the MAC OSX system voice saying “low” or “high.” Each scan run consisted of twelve 30 s blocks (6 per condition), and there were four scan runs per subject. The physical stimulus did not change across compared conditions, only the attentional state.

Each stream had a temporal pattern similar to Morse code making the task comparable to tuning into one of two competing tonal conversations at a time: patterns consisted of pseudorandomly intermixed long (300 ms) and short duration (75 ms) ramped tone bursts separated by blank intervals (75 ms). In each stream independently, the patterns were presented in a series of two-interval forced choice trials (2-IFC). During each trial, a randomly generated 5- or 6-element pattern (interval 1) was presented followed by a second 5- or 6-element pattern (interval 2) that was either identical to the first or a shuffled permutation of the first, and subjects made a “same” or “different” judgment within the cued attended stream only. The duration of interval 1 was up to 1350 ms (depending on the generated pattern) and the second interval started 2 s after onset of interval 1 (minimum interstimulus interval of 750 ms). Subjects had 1 s after the offset of interval 2 to enter their response by pressing one of two keys with the right hand. A new trial began every 4.7 s and each 30 s block had six trials. The sequence of the first interval, and whether the second interval was the same or different, was independently randomized in each stream every trial. Thus, subjects could perform the task only by attending to the cued stream. The starting condition was counterbalanced across subjects.

Before scanning, subjects participated in a brief training session (30 min). The patterns in the 2-IFC task could be either 5- or 6-element in length (6 being more difficult) and were for adjusted during training per subject to achieve performance that was well-above chance but not at ceiling. The number of elements used was then fixed per subject: 4 of the 6 subjects were given 6-element sequences. Percent correct performance was the same for attend high and attend low trials during fMRI scanning (See Results), indicating no difference in task difficulty across conditions.



**Figure 2.** *A*, Mean fMRI time courses during the dual-stream selective attention experiment (across all subjects and hemispheres,  $n = 12$ ) smoothed with a Gaussian (half-width = 8 s). Time courses were extracted from all voxels of primary auditory cortex labeled as preferring 250 Hz (light gray) and 4000 Hz (dark gray) in each subject and hemisphere based on individual subject tonotopic mappings. The responses of 4000 Hz-preferring voxels increased during attend high blocks and decreased during attend low blocks, and vice versa for 250 Hz-preferring voxels. *B*, Frequency attention. Bars show the mean difference in response between attend high and attend low blocks across all voxel bins in primary auditory cortex with all frequency preferences. *C*, Spatial attention. Bars show the mean difference in response between attend contralateral and attend ipsilateral blocks. *D*, Modulation in single-stream experiment. Bars show the mean difference in response between high and low blocks measured in separate scans in which the stimulus physically alternated between high-only and low-only streams. Note change in  $y$ -axis scale. Comparing the amplitudes in *B* and *C*, feature-selective attention outweighed effects of spatial attention by a factor of  $\sim 5$ . Comparing the amplitudes of *B* and *D*, frequency attention modulation was 18.6% as large as stimulus-driven modulation in 250 Hz voxels and 56.2% as large in 4000 Hz voxels, a robust modulatory effect. Error bars show SEM across all subjects and hemispheres,  $n = 12$ .

**Single-stream experiment.** Next, we asked: How does attending to one of the two concurrent frequency streams compare with hearing that frequency stream alone (i.e., complete disappearance of the ignored stimulus)? In the same subjects, we ran a second version of the experiment (single-stream experiment) in which the stimuli and task were the same as the first experiment except that the ignored stream was physically removed during each block. The stimulus physically alternated between a single attended high-frequency stream (4000 Hz) in one ear and a single attended low-frequency stream (250 Hz) in the other ear: “high versus low.” Hence, response modulations would include both stimulus-driven and attentional effects. Each subject performed four runs of the single-stream experiment alternately interleaved with the four runs of the dual-stream experiment. Single-stream runs were counterbalanced for ear-side and starting block in the same manner as the dual-stream experiment.

## Results

Figure 2*A* plots fMRI time courses recorded during the selective attention (dual-stream) experiment, across subjects and hemispheres ( $n = 12$ ). Specifically, time courses were extracted from all volumetric primary auditory cortex voxels (hA1 and hR combined) labeled as having best frequencies of 250 Hz in red and 4000 Hz in blue on a per-subject, per-hemisphere basis based on the individual's own tonotopic mapping (mean number of 250 Hz-tuned voxels:  $320 \pm 140$  SD; 4000 Hz-tuned voxels:  $115 \pm 55$  SD across subjects and hemispheres). Each plotted time courses is an average of the 12 extracted time courses (one per-subject, per-hemisphere). As can be appreciated by eye, the responses of 4000 Hz-tuned voxels increased during the attend high condition and decreased during the attend low condition, while in 250 Hz-tuned voxels, the opposite pattern of modulation was seen. The stimulus itself and task difficulty did not change across blocks



(task performance: attend low blocks =  $89.3 \pm 9.1\%$  SD; attend high blocks =  $87.3 \pm 7.4\%$  SD across subjects,  $p = 0.59$ , paired  $t$  test), and thus we attribute this modulation in primary auditory cortex to frequency-selective attention.

Next, we show response modulation not only within the 250 and 4000 Hz best-frequency voxels, but across all primary auditory cortex voxels with all frequency preferences (from 88 to 8000 Hz in half-octave bins). The bars in Figure 2B indicate the mean difference between attend high blocks and attend low blocks (adjusted 4 s for hemodynamic delay) across all voxel bins (means and SE bars computed over individual responses per-subject, per-hemisphere,  $n = 12$ ). Overall, the responses of frequency-tuned units were enhanced by attention to preferred versus non-preferred frequencies. Response modulations were highly significant in 250 Hz and 4000 Hz voxels ( $p < 0.0005$  and  $p < 0.00005$ , respectively,  $t$  test) and Figure 2B shows the overall tuning profile of the frequency attention-effect. Data from hA1 and hR voxels from both hemispheres are combined here since the pattern of modulation was qualitatively similar and individually significant when analyzed separately in hA1 and hR voxels (A1,  $p < 0.001$  and  $p < 0.01$ ; R,  $p < 0.001$  and  $p < 0.001$  for 250 Hz and 4000 Hz voxels, respectively) and in left and right hemispheres (LH,  $p < 0.005$  and  $p < 0.01$ ; RH,  $p < 0.01$  and  $p < 0.005$  for 250 Hz and 4000 Hz voxels). Further, to verify the reliability of our manual ROI selection, we subsequently had three experimenters (S.D.C., W.V.D.Z., and M.S.) independently draw the primary auditory cortex ROI: inter-rater overlap was high (RH, 0.87; LH, 0.84; Dice coefficient), and the pattern of results was unchanged when we reanalyzed only those voxels which overlapped all three selections ( $p < 0.0001$  and  $p < 0.00001$  and for 250 Hz and 4000 Hz voxels, respectively).

The data plotted so far compare responses to attend high versus attend low conditions, regardless of stimulus side. Next, we look at the effect of attending to contralateral versus ipsilateral sides, regardless of stimulus frequency. Figure 1C plots the mean difference in response between attend contralateral and attend ipsilateral blocks across all voxel bins. A general response increase for attending the contralateral side is observed (spatial attention effect). The profiles of frequency-attention and spatial-attention effects are different: the effects of frequency-selective attention (Fig. 2B) are largest in voxels near the frequencies used (250 and 4000 Hz) and taper off gradually in voxels tuned to neighboring frequencies, while the effects of spatial attention (Fig. 2C) are similar in magnitude across voxels with all frequency preferences. The effects of frequency-selective attention outweighed the more modest effects of spatial-selective attention by a factor of 4.9 in 250 Hz voxels and a factor 5.8 in 4000 Hz voxels.

Finally, we address the question: how does focusing attention onto one of two competing stimuli compare with making the ignored stimulus physically disappear? In separate interleaved experimental runs (single-stream experiment), the stimulus physically alternated between a single attended high-frequency stream in one ear and a single attended low-frequency stream in the other ear (high vs low). Figure 2D plots the mean difference in response between high and low blocks in the single-stream experiment across all voxel bins. Task scores indicated no difference in difficulty across blocks (low =  $94.2 \pm 3.8\%$  SD, high =  $93.9 \pm 2.6\%$  SD across subjects,  $p = 0.8$ , paired  $t$  test). Next, we compare the modulation amplitudes of Figure 2B (in which only attentional state alternated) to the modulation amplitudes of Figure 2D (in which the physical stimulus alternated). In 250 Hz voxels, we see that attentional modulation was 18.6% as large as modulation due to physically alternating the stimulus; and in 4000 Hz

voxels it was 56.2% as large. The difference in percentages between 250 Hz and 4000 Hz voxels reflects the denominator: the modulation due to physically alternating the stimulus was stronger in low-frequency voxels compared with high-frequency voxels, consistent with previous reports of weaker BOLD responses to high-frequency stimuli for reasons not fully understood (Langers and van Dijk, 2012). In either case, frequency-selective attention can be regarded as a powerful modulatory effect.

## Discussion

We demonstrated that neural activity within human primary auditory cortex (hA1 and hR) is strongly and dynamically modulated by attention to preferred versus nonpreferred sound frequencies. These effects of frequency-attention outweighed more modest effects of spatial-attention by a factor of  $\sim 5$  and were up to 56% as large as when physically removing the competing stimulus. The frequency-attention effect was largest in voxels near the specific attended frequencies (250 Hz and 4000 Hz) and tapered off gradually in voxels tuned to neighboring frequencies. The results suggest that, like a radio, primary auditory cortex can tune into attended frequency channels and can rapidly switch channels to meet task demands.

These results are consistent with the previous human fMRI study by Paltoglou et al. (2009) that showed frequency-specific attentional modulation of auditory cortex, although with a less detailed frequency mapping. Interestingly, Oh et al. (2012) demonstrated frequency-specific modulation of auditory cortex during imagery of low- versus high-frequency tones, which may rely on related mechanisms of top-down modulation. Our study adds to the previous findings by performing high-resolution, fine-scaled frequency mappings which allow us to (1) unambiguously identify primary auditory cortical fields, and (2) characterize the tuning of attentional effects as a function of frequency preference. Our findings are also novel in that the experimental design allowed comparison of frequency-attention effects to spatial attention and stimulus-driven effects.

The degree of attentional modulation observed in a region likely depends on the extent to which the underlying neurons encode the features of the attended target, and it may not be surprising that we observed greater modulation to shifts in attended frequency compared with shifts in attended location. Across many species, primary auditory cortex contains a fine-grained representation of sound frequency and is organized tonotopically (Bitterman et al., 2008; Bartlett et al., 2011; Guo et al., 2012), but spatial tuning is notably broad and a cortical topographic organization has not been found (Recanzone, 2000). Unilaterally presented sounds are known to induce a significant bilateral fMRI response in human auditory cortex (van der Zwaag et al., 2011). However modestly, we did observe spatial attention effects in A1 and R; and previous studies have shown spatially-driven attentional modulation in auditory cortex (Rinne et al., 2008, 2012). It is possible that spatial attentional modulation would be greater with a task that required more use of spatial information.

Our results (both dual stream and single stream) showed a fairly broad frequency-tuning at the voxel level, larger than what is expected of some individual neurons. This broad tuning could be related to a population level mixture of narrowly and broadly tuned neurons ( $\sim 10^4$ – $10^5$  neurons per cubic mm in cortex). In rat A1 the precise tonotopic organization of middle cortical layers is degraded in the superficial and deep cortical layers, where many irregularly tuned neurons are found (Guo et al., 2012). Additionally, broadened frequency-tuning is expected at high

stimulus sound intensities (Tanji et al., 2010; Guo et al., 2012), an effect which originates at the basilar membrane. However, it should be noted that sharp frequency tuning, on the order of 1/12th of an octave, was found to suprathreshold sound stimuli in a large proportion of A1 single neurons in alert humans (intracranial depth electrodes, Bitterman et al., 2008), and in awake-behaving marmosets (Bartlett et al., 2011). Thus we expect that some component of our population BOLD response in humans arises from sharply frequency-tuned neurons.

### Comparison to single-neuron studies of auditory attention

The observed attentional modulation in the BOLD response could reflect both neural enhancement and suppression, and recent findings from single-neuron recordings in animals emphasize the role of both in the modulation of A1 receptive fields. For example, in ferret A1, individual frequency-tuning profiles were rapidly reshaped when animals attended to target tones amid distractor sounds (Fritz et al., 2003, 2005, 2007; Atiani et al., 2009; David et al., 2012). In many cases, neurons tuned near the target frequency showed enhanced responsiveness to best frequency and those tuned to background frequencies showed suppression, but, interestingly, target frequency suppression could be evoked under different behavioral contexts (David et al., 2012). In rat A1, neurons showed enhanced responses during attention to target tones matching the neuron's best frequency (Jaramillo and Zador, 2011), but also showed broad suppression during performance of an auditory task compared with passive listening (Otazu et al., 2009). Thus, it seems that A1 uses multiple strategies, not limited to target response enhancement, to sharpen the representation of attended stimuli relative to background. It is not straightforward to relate these single-neuron findings in animals to BOLD population results in humans, except to say that a combination of attention-related excitatory and inhibitory mechanisms could contribute to the observed BOLD modulation. One difference between our study and the single-neuron studies cited here is that the effects in animals followed many weeks of specific task training and could persist minutes to hours after task completion (Fritz et al., 2003, 2007). Thus, the effects in animals possibly depend upon long-term learning mechanisms, in addition to short-term flexible attentional mechanisms. Our study demonstrates dynamic and transient (target shifting every 30 s) attentional modulation of A1 and R using a task that required only limited training in humans.

### Comparison to feature-based attention in visual cortex

Our findings are broadly consistent with 'feature-based' models of attention, from the visual cortex literature, which propose that responses are enhanced in neurons whose feature-selectivity matches the current attentional focus (Treue and Martinez-Trujillo, 1999; Saenz et al., 2002; Maunsell and Treue, 2006). In visual cortex, feature-based attention has been shown to modulate both stimulus-evoked responses and spontaneous baseline activity in the absence of a stimulus (macaque single unit: Luck et al., 1997; Reynolds et al., 1999; human fMRI: Serences and Boynton, 2007). Thus, feature-based attention could serve both to strengthen the neuronal representation of an attended target and/or increase the detectability of an anticipated target if its features are known in advance. Likewise it is possible that attention to a sound frequency could modulate the baseline activity of auditory cortex neurons in the absence of a stimulus.

### Broader significance

Frequency is one featural cue out of several, including position, trajectory, timbre, intensity, and temporal cues, that likely contribute to speech selection (Zion Golumbic et al., 2012). Responses to attended speech patterns are enhanced and responses to unattended speech patterns suppressed at higher levels of auditory cortex (Kerlin et al., 2010; Mesgarani and Chang, 2012) and age-related deficits in speech comprehension in noise are linked to impaired attentional mechanisms in older adults (Passow et al., 2012). Spectral filtering by attention may be an important function of the primary auditory cortex, contributing to downstream selection of spectrally complex auditory streams such as speech.

### References

- Ahveninen J, Hämäläinen M, Jääskeläinen IP, Ahlfors SP, Huang S, Lin FH, Raj T, Sams M, Vasios CE, Belliveau JW (2011) Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proc Natl Acad Sci U S A* 108:4182–4187. [CrossRef Medline](#)
- Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA (2009) Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive field. *Neuron* 61:467–480. [CrossRef Medline](#)
- Bartlett EL, Sadagopan S, Wang X (2011) Fine frequency tuning in monkey auditory cortex and thalamus. *J Neurophysiol* 106:849–859. [CrossRef Medline](#)
- Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J Neurosci* 27:9252–9261. [CrossRef Medline](#)
- Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature* 451:197–201. [CrossRef Medline](#)
- Cherry EC (1953) Some experiments on the recognition of speech, with one and two ears. *J Acoust Soc Am* 25:975–979. [CrossRef](#)
- Da Costa S, van der Zwaag W, Marques JP, Frackowiak RSJ, Clarke S, Saenz M (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci* 31:14067–14075. [CrossRef Medline](#)
- David SV, Fritz JB, Shamma SA (2012) Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc Natl Acad Sci U S A* 109:2144–2149. [CrossRef Medline](#)
- Engel SA (2012) The development and use of phase-encoded functional MRI designs. *Neuroimage* 62:1195–1200. [CrossRef Medline](#)
- Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R (2003) Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* 40:859–869. [CrossRef Medline](#)
- Fritz JB, Elhilali M, Shamma SA (2005) Different dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J Neurosci* 25:7623–7635. [CrossRef Medline](#)
- Fritz JB, Elhilali M, Shamma SA (2007) Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J Neurophysiol* 98:2337–23346. [CrossRef Medline](#)
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223. [CrossRef Medline](#)
- Fujiwara N, Nagamine T, Imai M, Tanaka T, Shibasaki H (1998) Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. *Brain Res Cogn Brain Res* 7:99–109. [CrossRef Medline](#)
- Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG, Polley DB (2012) Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J Neurosci* 32:9159–9172. [CrossRef Medline](#)
- Hackett TA (2011) Information flow in the auditory cortical network. *Hear Res* 271:133–146. [CrossRef Medline](#)
- Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. *Science* 182:177–180. [CrossRef Medline](#)
- Humphries C, Liebenthal E, Binder JR (2010) Tonotopic organization of human auditory cortex. *Neuroimage* 50:1202–1211. [CrossRef Medline](#)

- Jäncke L, Mirzazade S, Shah NJ (1999) Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266:125–128. [CrossRef Medline](#)
- Jaramillo S, Zador AM (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci* 14:246–251. [CrossRef Medline](#)
- Kerlin JR, Shahin AJ, Miller LM (2010) Attentional gain control of ongoing cortical speech representations in a “cocktail party.” *J Neurosci* 30:620–628. [CrossRef](#)
- Langers DR, van Dijk P (2012) Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb Cortex* 22:2024–2038. [CrossRef Medline](#)
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77:24–42. [Medline](#)
- Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R (2010) MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage* 49:1271–1281. [CrossRef Medline](#)
- Maunsell JHR, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322. [CrossRef Medline](#)
- Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485:233–236. [CrossRef Medline](#)
- Oh JH, Kwon JH, Yang PS, Jeong J (2012) Auditory imagery modulates frequency-specific areas in the human auditory cortex. *J Cogn Neurosci*. Advance online publication. Retrieved September 1, 2012. doi:10.1162/jocn\_a\_00280. [CrossRef Medline](#)
- Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. *Nat Neurosci* 12:646–654. [CrossRef Medline](#)
- Paltoglou AE, Sumner CJ, Hall DA (2009) Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. *Hear Res* 257:106–118. [CrossRef Medline](#)
- Passow S, Westerhausen R, Wartenburger I, Hugdahl K, Heekeren HR, Lindenberger U, Li SC (2012) Human aging compromises attentional control of auditory perception. *Psychol Aging* 27:99–105. [CrossRef Medline](#)
- Petkov CI, Kang X, Alho K, Bertrand O, Yund EW, Woods DL (2004) Attentional modulation of human auditory cortex. *Nat Neurosci* 7:658–663. [CrossRef Medline](#)
- Recanzone GH (2000) Spatial processing in the auditory cortex of the macaque monkey. *Proc Natl Acad Sci U S A* 97:11829–11835. [CrossRef Medline](#)
- Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19:1736–1753. [Medline](#)
- Rinne T, Balk MH, Koistinen S, Autti T, Alho K, Sams M (2008) Auditory selective attention modulates activations of human inferior colliculus. *J Neurophysiol* 100:3323–3327. [CrossRef](#)
- Rinne T, Koistinen S, Talja S, Wikman P, Salonen O (2012) Task-dependent activations of human auditory cortex during spatial discrimination and spatial memory tasks. *Neuroimage* 59:4126–4131. [CrossRef Medline](#)
- Rivier F, Clarke S (1997) Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage* 6:288–304. [CrossRef Medline](#)
- Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. *Nat Neurosci* 5:631–632. [CrossRef Medline](#)
- Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301–312. [CrossRef Medline](#)
- Striem-Amit E, Hertz U, Amedi A (2011) Extensive cochleotopic mapping of human auditory cortical fields obtained with phase-encoded fMRI. *PLoS One* 6:e17832. [CrossRef Medline](#)
- Tanji K, Leopold DA, Ye FQ, Zhu C, Malloy M, Saunders RC, Mishkin M (2010) Effect of sound intensity on tonotopic fMRI maps in the unanesthetized monkey. *Neuroimage* 49:150–157. [CrossRef Medline](#)
- Treue S, Martinez-Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 299:575–579.
- van der Zwaag W, Francis S, Head K, Peters A, Gowland P, Morris P, Bowtell R (2009) fMRI at 1.5, 3, and 7 T: characterizing BOLD signal changes. *Neuroimage* 47:1425–1434. [CrossRef Medline](#)
- van der Zwaag W, Gentile G, Gruetter R, Spierer L, Clarke S (2011) Where sound position influences sound object representations: a 7-T fMRI study. *Neuroimage* 54:1803–1811. [CrossRef Medline](#)
- Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci U S A* 90:8722–8726. [CrossRef Medline](#)
- Woods DL, Hillyard SA, Hansen JC (1984) Event-related brain potentials reveal similar mechanisms during selective listening and shadowing. *J Exp Psychol Hum Percept Perform* 10:761–777. [CrossRef Medline](#)
- Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X (2009) Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One* 4:e5183. [CrossRef Medline](#)
- Woods DL, Herron TJ, Cate AD, Yund EW, Stecker GC, Rinne T, Kang X (2010) Functional properties of human auditory cortical fields. *Front Syst Neurosci* 4:155. [Medline](#)
- Zion Golumbic EM, Poeppel D, Schroeder CE (2012) Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain Lang* 122:151–161. [CrossRef Medline](#)