JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Systems/Circuits

Extensive tonotopic mapping across auditory cortex is recapitulated by spectrally-directed attention, and systematically related to cortical myeloarchitecture

Frederic K. Dick[1,2,3], Matt I. Lehet[4,5], Martina F. Callaghan[7], Tim A. Keller[4,6], Martin I. Sereno[8] and Lori L. Holt[4,5]

[1]*Department of Psychological Sciences, Birkbeck College, University of London, London, WC1E 7HX*
[2]*Birkbeck-UCL Centre for Neuroimaging, London, WC1H 0AP*
[3]*Department of Experimental Psychology, University College London, London, WC1H 0AP*
[4]*Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213*
[5]*Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213*
[6]*Scientific Imaging & Brain Research Center, Carnegie Mellon University, Pittsburgh, PA 15213*
[7]*Wellcome Trust Center for Neuroimaging, Institute of Neurology, University College London, London, WC1N 3BG*
[8]*Department of Psychology, San Diego State University, San Diego, CA San Diego, CA 92182-4611*

DOI: 10.1523/JNEUROSCI.1436-17.2017

Received: 24 May 2017

Revised: 4 October 2017

Accepted: 6 October 2017

Published: 6 November 2017

Author contributions: F.D. and L.L.H. designed research; F.D., M.I.L., T.K., and L.L.H. performed research; F.D., M.I.L., and L.L.H. analyzed data; F.D., M.F.C., M.S., and L.L.H. wrote the paper; M.F.C. and M.S. contributed unpublished reagents/analytic tools.

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

Research supported by Rothberg Research Award in Human Brain Imaging of Carnegie Mellon University. M.L. was supported by NIH T90DA022761 and T32GM081760. Thanks to Scott Kurdilla and Debbie Viszlay (SIBR) for imaging support, Antoine Lutti and Nikolaus Weiskopf for physics expertise and generosity with porting the MPM protocol to the SIBR Verio, and the developers of FSL, AFNI, and FreeSurfer for their scientific work and software. Finally, we are grateful to Marlene Behrmann, Jenny Bizley, Maria Chait, Tim Griffiths, Jen Linden, Catherine Perrodin, Chris Petkov, Lars Riecke, Sam Schwarzkopf, Jeremy Skipper, Ediz Sohoglu, Adam Tierney, and three anonymous reviewers of a previous version of the manuscript for extremely useful suggestions and feedback. The Wellcome Centre for Human Neuroimaging is supported by core funding from Wellcome [203147/Z/16/Z].

Corresponding Authors: Frederic K. Dick, Birkbeck/UCL Centre for NeuroImaging, 26 Bedford Way, London, WC1H 0AP, UK +44 20 7079 0815, f.dick@bbk.ac.uk; Lori L. Holt, Department of Psychology, Baker Hall, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, +1 (412) 268-4964, loriholt@cmu.edu

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.1436-17.2017

Extensive tonotopic mapping across auditory cortex is recapitulated by spectrally-directed attention, and systematically related to cortical myeloarchitecture

Frederic K. Dick [1,2,3] (ORCID ID 0000-0002-2933-3912)
Matt I. Lehet [4,5] (ORCID 0000-0002-7601-9360)
Martina F. Callaghan [7] (ORCID ID 0000-0003-0374-1659)
Tim A. Keller [4,6] (ORCID ID 0000-0002-2345-8626)
Martin I. Sereno [8] (ORCID ID 0000-0002-7598-7829)
Lori L. Holt [4,5] (ORCID ID 0000-0002-8732-4977)

[1] Department of Psychological Sciences, Birkbeck College, University of London, London, WC1E 7HX
[2] Birkbeck-UCL Centre for Neuroimaging, London, WC1H 0AP
[3] Department of Experimental Psychology, University College London, London, WC1H 0AP
[4] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213
[5] Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213
[6] Scientific Imaging & Brain Research Center, Carnegie Mellon University, Pittsburgh, PA 15213
[7] Wellcome Trust Center for Neuroimaging, Institute of Neurology, University College London, London, WC1N 3BG
[8] Department of Psychology, San Diego State University, San Diego, CA San Diego, CA 92182-4611

Corresponding Authors:

Frederic K. Dick, Birkbeck/UCL Centre for NeuroImaging, 26 Bedford Way, London, WC1H 0AP, UK +44 20 7079 0815, f.dick@bbk.ac.uk

Lori L. Holt, Department of Psychology, Baker Hall, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, +1 (412) 268-4964, loriholt@cmu.edu

Abbreviated Title: Spectrally-directed attention maps in auditory cortex
Number of Pages: 49
Number of Figures: 8
Number of Words for Abstract: 228
Number of Words for Introduction: 655
Number of Words for Discussion: 1402
Conflict of Interest: The authors declare no competing financial interests.

Acknowledgements: Research supported by Rothberg Research Award in Human Brain Imaging of Carnegie Mellon University. M.L. was supported by NIH T90DA022761 and T32GM081760. Thanks to Scott Kurdilla and Debbie Viszlay (SIBR) for imaging support, Antoine Lutti and Nikolaus Weiskopf for physics expertise and generosity with porting the MPM protocol to the SIBR Verio, and the developers of FSL, AFNI, and FreeSurfer for their scientific work and software. Finally, we are grateful to Marlene Behrmann, Jenny Bizley, Maria Chait, Tim Griffiths, Jen Linden, Catherine Perrodin, Chris Petkov, Lars Riecke, Sam Schwarzkopf, Jeremy Skipper, Ediz Sohoglu, Adam Tierney, and three anonymous reviewers of a previous version of the manuscript for extremely useful suggestions and feedback. The Wellcome Centre for Human Neuroimaging is supported by core funding from Wellcome [203147/Z/16/Z].

Abstract

Auditory selective attention is vital in natural soundscapes. But, it is unclear how attentional focus on the primary dimension of auditory representation - acoustic frequency - might modulate basic auditory functional topography during active listening. In contrast to visual selective attention, which is supported by motor-mediated optimization of input across saccades and pupil dilation, the primate auditory system has fewer means of differentially sampling the world. This makes spectrally-directed endogenous attention a particularly crucial aspect of auditory attention. Using a novel functional paradigm combined with quantitative MRI, we establish in male and female listeners that human frequency-band-selective attention drives activation in both myeloarchitectonically-estimated auditory core, and across the majority of tonotopically-mapped non-primary auditory cortex. The attentionally-driven best-frequency maps show strong concordance with sensory-driven maps in the same subjects across much of the temporal plane, with poor concordance in areas outside traditional auditory cortex. There is significantly greater activation across most of auditory cortex when best frequency is attended, versus ignored; the same regions do not show this enhancement when attending to the least-preferred frequency band. Finally, the results demonstrate that there is spatial correspondence between the degree of myelination and the strength of the tonotopic signal across a number of regions in auditory cortex. Strong frequency preferences across tonotopically-mapped auditory cortex spatially correlate with R_1-estimated myeloarchitecture, indicating shared functional and anatomical organization that may underlie intrinsic auditory regionalization.

77 **Significance**

78

79 Perception is an active process especially sensitive to attentional state. Listeners direct auditory
80 attention to track a violin's melody within an ensemble performance, or to follow a voice in a
81 crowded cafe. Although diverse pathologies reduce quality of life by impacting such spectrally-
82 directed auditory attention, its neurobiological bases are unclear. We demonstrate that human
83 primary and non-primary auditory cortical activation is modulated by spectrally-directed attention
84 in a manner that recapitulates its tonotopic sensory organization. Further, the graded activation
85 profiles evoked by single frequency bands are correlated with attentionally-driven activation
86 when these bands are presented in complex soundscapes. Finally, we observe a strong
87 concordance in the degree of cortical myelination and the strength of tonotopic activation across
88 several auditory cortical regions.

89

Introduction

Listeners shift attention across multiple simultaneously-present acoustic dimensions to home in on those that are diagnostic in guiding behavior (Idemaru and Holt, 2011; Herrmann and Henry, 2013; Shamma and Fritz, 2014). In nonhuman animal studies task-based spectral attention adaptively modulates auditory neurons' spectrotemporal response fields (Fritz et al., 2010). Human neuroimaging results reveal that attention to streams of high- versus low-frequency acoustic input can modulate activity in tonotopically-defined regions (Paltoglou et al., 2009), as can imagery of higher versus lower frequencies (Oh et al., 2013). In and directly around Heschl's gyrus, there are strong frequency-band-specific attentional effects to high and low pure-tone streams presented to opposite ears (Da Costa et al., 2013) and a shared topography of sensory and attentionally-driven responses (Riecke et al., 2016). These results establish that endogenous attention directed across acoustic frequency, the primary axis of auditory representation, can modulate human cortical activity in a tonotopic manner around Heschl's gyrus. But, there remain important unanswered questions about the neurobiological basis of human spectrally-directed attention.

First, does the topography of attention to different frequency bands recapitulate tonotopic organization in human primary auditory cortex? Non-human animal physiology establishes spectrally-directed attention in myelo- and cyto-architectonically-defined primary areas in 'auditory core' (Fritz et al., 2007b; Shamma and Fritz, 2014). However, although two recent neuroimaging studies have shown strong similarities between stimulus- and attentionally-driven tonotopic organization in and directly around Heschl's gyrus (Da Costa et al., 2013; Riecke et al., 2016), it has not yet been possible to unambiguously localize this effect to human auditory core. Here, we use high-resolution quantitative MRI (Pierpaoli, 2010) to estimate myelo-

architectonically-defined auditory core, and demonstrate that spectrally-directed attention modulates its activation in a tonotopically-organized manner.

Second, is attentionally-driven tonotopic organization present outside of auditory core? In humans, Riecke et al. (2016) found no significant evidence for tonotopically organized effects of spectral attention outside of early auditory areas, but did show that the information content of non-primary cortical frequency representations was sufficient for above-chance decoding of listeners' frequency-selective attentional focus. The lack of attentionally-driven tonotopic maps contrasts with the finding that most non-primary cortical visual areas exhibit strong retinotopically-specific attentional effects (Saygin and Sereno, 2008). Using intensive data collection (>7000 functional volumes per subject) we present evidence for widespread, tonotopically-organized modulation by spectral attention across much of auditory cortex, with individual differences in individual participants' tonotopic maps reproduced in attentionally-driven maps.

Third, what is the effect of frequency-selective attention being directed to a voxel's non-preferred frequency band? Detailed fMRI studies of stimulus-driven frequency response functions (Schönwiesner and Zatorre, 2009; Moerel et al., 2013) have shown graded and multi-peaked frequency responses across human auditory cortex. However, it is unclear whether these more complex patterns are recapitulated by attention to a given frequency band. In the context of three distinct frequency bands, (Riecke et al., 2016) found that attentional filters appeared to be bandpass in and around Heschl's gyrus. Here, using a five-frequency-band paradigm, we establish that graded response profiles evoked by single frequency bands are strongly associated with attentionally-driven response profiles to those frequencies across much of auditory cortex. We also show that a systematic topography of 'dis-preferred' frequency can be driven by attention, and establish the regionalization of spectral attentional effects relative to prior studies of crossmodal auditory attention (Petkov et al., 2004).

Finally, is there spatial correspondence between auditory cortical anatomy, as measured by the local change in R_1-estimated myelination, and fMRI-assessed strength of relative frequency selectivity? Post-mortem Gallyas staining to establish human myeloarchitecture reveals considerable variability in auditory cortical myelination that is associated with MRI signal change in the same brain (Wallace et al., 2016). Likewise, variation in cortical myelination estimated using T1-weighted/T2-weighted ratio approaches also appears to correspond spatially with some functional variation in the superior temporal lobe (Glasser et al., 2016). Here, we demonstrate that there is spatial concordance between the degree of myelination and the amplitude of the frequency-selective tonotopic signal across several regions in auditory cortex.

Methods

Experiment Overview. We used a novel paradigm in which listeners direct attention to a series of four-tone 'mini-sequences' that fall within one of five possible spectral bands, without any spatial cues. The task is to monitor for temporally-adjacent mini-sequence repeats within the attended band. Inasmuch as this places a very high demand on encoding and integrating spectral sequences within a delimited frequency range, we expect it to be especially effective in evoking strong responses in non-primary auditory cortical areas. The goal is to address where, specifically, in the auditory system spectral gain from attention is evident, and akin to long-standing work in vision (Kastner and Ungerleider, 2000), to delineate the topographic maps across which attentional modulation is apparent.

The target mini-sequences were embedded in either an informationally-sparse or informationally-dense acoustic scene (Figure (Fig.) 1). Streams of four-tone mini-sequences were presented in either a single band ('tonotopy', Fig. 1a), or accompanied by mini-sequences in a 'distractor' frequency band, the center frequency of which varied in the frequency distance

from the attended band across blocks (‘attention-tonotopy’, abbreviated 'attn-tono', Fig. 1b). A verbal cue directed listeners’ attention to a specific frequency band, within which listeners monitored four-tone mini-sequences for repeats; the ‘distractor’ band in attn-tono blocks also contained repeats. Using a discretized version of a phase-encoded fMRI design (Sereno et al., 1995; Schwarzkopf et al., 2011; Rao and Talavage, 2005; Herdener et al., 2013; Langers et al., 2014), the cued frequency band stepped up or down in orderly steps across the acoustic spectrum across a 64-sec cycle [Fig. 1c]. Phase-encoded tonotopic designs benefit from the power and robustness of the 'travelling wave' method for topographic cortical mapping of smoothly varying representations (Engel, 2012); the discretized (blocked) version we use here allows use of the verbal cue, and has the advantage of being able to be analyzed using both Fourier and regression approaches. This allowed us to include an additional, ‘randomized’ attn-tono condition that contributed both as a control condition in Fourier analyses, and also as an additional attn-tono run in regression analyses (see Fig. 1d). The tone stimuli from this condition were *identical* to the ‘stepped’ attn-tono condition, but the order of the verbal cues directing listeners’ attention to a specific frequency band was scrambled in their assignment to blocks. This preserved the acoustics across conditions, but eliminated the consistent ‘stepping’ of attention through the frequency spectrum in the randomized condition thereby destroying the consistent phase-lag associated with specific frequency bands that support Fourier analyses [schematized in Fig. 1e, and see below]. Thus, to the extent that there are attentionally-driven frequency-selective maps in auditory cortex we expect tonotopically-organized attentional maps to be apparent in the 'stepped,' but not the 'randomized' attn-tono conditions under Fourier analyses. In contrast, regression analyses include a model of attention, allowing ‘stepped’ and ‘randomized’ attn-tono conditions to be pooled to investigate the impact of attention on cortical activation. Across both Fourier and regression analyses, the ‘stepped’ attn-tono conditions were collapsed across runs for which the cued frequency band stepped up in frequency and those

189 that stepped down; inclusion of each simply balanced the directional movement of attention
190 through the acoustic spectrum across the experiment.
191
192 In summary, in the *attn-tono* conditions, attention alone was available to differentially drive
193 responses to an approximately constant acoustic input, whereas in the *tonotopy* condition,
194 responses were driven by spectrally-selective stimuli as well as by attention.
195
196 We analyzed mapping data using Fourier methods with individual and surface-based group
197 analysis methods, as described previously (Sereno et al., 1995; Hagler et al., 2006; 2007). With
198 this approach, voxels preferentially responding to a certain phase in a 'stepped' stimulus cycle
199 are defined as those that have a significantly higher signal amplitude at this stimulus temporal
200 frequency (meaning the slow frequency of the repeat of the spectral ramp) than at the average
201 of other 'noise' frequencies (see Fig. 1e). Significant signal *phases* (a particular position in the
202 cycle) are then mapped to a color scale to indicate the voxel's 'best frequency' and signal
203 amplitude is mapped to the voxel's color saturation (Fig. 1f). We time-reversed runs stepping
204 down in frequency and averaged them with runs stepping up in frequency (Sereno et al., 1995;
205 Talavage et al., 2004; Dick et al., 2012; Ahveninen et al., 2016). Cross-subject averaging of
206 phase-encoded mapping data was performed using a method described previously (Hagler et
207 al., 2007) in which the real and imaginary components of the signal with respect to the stepped
208 cycle were sampled to the cortical surface and then averaged across subjects, preserving any
209 phase information that was coherent over subjects.
210
211 Using previously established methods (Dick et al., 2012; Sereno et al., 2013; Lutti et al., 2014)
212 see also (Glasser et al., 2016), we used high-resolution quantitative multi-parameter mapping to
213 generate maps of estimated cortical myelination based on longitudinal relaxation times
214 (quantitative T_1). Recent work by multiple labs supports the hypothesis that T_1 relaxation is

reliably associated with quantitative differences in myelination in white matter and cortex (Sereno et al., 2013; Callaghan et al., 2014; Stüber et al., 2014; Dinse et al., 2015; Tardif et al., 2015; Turner, 2015; Tardif et al., 2016). Here, we calculated each subject's R_1 ($1/T_1$) values, where the greater the R_1, the higher the inferred myelin content. These R_1 values were resampled onto his or her surface at a cortical depth fraction of 0.5, and also averaged across individuals using sulcus-aligned cortical-surface-based procedures (see below for further details).

Participants. Eight adults (aged 23-45 y, mean 28 y; 6 female) participated; none reported a history of neurological disease or communication disorders. All had some childhood and/or adult musical training (one had a music degree), and had previous experience with longer scanning sessions. While musical training seemed to facilitate learning the experimental task, subsequent behavioral studies in the lab have shown that musically-naive subjects can also attain excellent performance with similar levels of training on this and even more demanding related tasks.

Stimuli and Design. Stimuli were created using custom code in Matlab version 2015a (Mathworks, Inc.) and SoX version 14.4.2 (sourceforge.net). The basic stimulus unit was a four-tone *mini-sequence* (140 ms sine-wave tones including 10 ms linear amplitude ramp), with each tone drawn with replacement from a seven-semitone, band-delimited pool centered around one of five frequencies (300, 566, 1068, 2016, 3805 Hz; Fig. 1a). Fourteen mini-sequences formed a *block* (mean inter-sequence silent interval 240 ms, SD 10 ms). Each block contained one to three mini-sequence repeats (1:2:1 ratio of 1, 2, and 3 repeats). When there was more than one repeat per block, mini-sequence repeat pairs were separated by at least one intervening mini-sequence. Each block began with a verbal prompt: '*hear', 'high',* or '*low';* loudness-equalized Mac 'Victoria' voice, mean duration 506 ms (SD 36 ms), padded with silence to 800 ms total

duration. This prompt was followed by 800 ms silent gap (tonotopy) or tone-cue (attn-tono), then the 14 mini-sequences (11.2s in total), for a total block duration of 12.8s.

The task was to detect mini-sequence repeats in the attended frequency band (i.e., a 1-back task). In the tonotopy condition, mini-sequences were confined to a single frequency band preceded by the neutral verbal prompt '*hear*' (Fig. 1a). In two of the four single-band tonotopy runs, block center frequency was stepped from low-to-high over a 64-second cycle with 8 cycles/run; step direction was reversed (high-to-low) for the other two runs. This is a 'discrete' version of phase-encoded designs commonly used in visual, somatosensory, and auditory mapping studies (Engel et al., 1994; Sereno et al., 1995; Da Costa et al., 2011; Dick et al., 2012; Langers and van Dijk, 2012; Langers et al., 2014; Saenz and Langers, 2014).

The attn-tono condition had the exact mini-sequence patterns from the tonotopy blocks, but there also were simultaneous, competing mini-sequences in a distinct frequency band with a center frequency at least 14 semitones apart (Fig. 1b; 300 vs. 1068 Hz; 300 vs. 2016 Hz; 300 vs. 3805 Hz; 566 vs. 2016 Hz; 566 vs. 3805 Hz; 1068 vs. 3805 Hz; not all center frequencies were paired due to the 14-semitone constraint). The verbal prompt (*high* or *low*) initiating each attn-tono block signaled participants to perform the 1-back task on either the higher or lower frequency band. Immediately after the verbal prompt, a randomly-ordered pair of sine-wave tones cued the center frequencies of the upcoming block (140 ms tones including 8 ms linear on/off amplitude ramp; 26 ms inter-tone silence, tone pair followed by 494 ms silence, total duration 800 ms). Crucially, there were mini-sequence repeats even in the unattended band to assure that attention directed to the task was endogenously driven rather than being attracted by stimulus repetition effects (Barascud et al., 2016).

There were two attn-tono conditions: 'stepped' and 'randomized'. Analogous to single-band tonotopy runs, in stepped attn-tono runs the verbally-cued frequency band implicitly stepped up (2 runs) or down (2 runs) in frequency over a 64-sec *cycle* (Fig. 1c). This cued iterative stepping through the frequency spectrum facilitates transfer of attention to each frequency band (as in traditional phase-encoded designs) and supports Fourier approaches to analysis (Fig. 1e). Each randomized attn-tono run was acoustically identical to a stepped run, but the verbal prompt was manipulated so that there was no systematic, stepped organization of mini-sequence center frequencies through the spectrum (Fig. 1d). For this condition, frequency bands were cued at inconsistent phase lags within the 8 cycles/run, thereby phase-canceling any periodic attentional response; this is schematized in Fig 1e. This randomized-order control is important, as there is a small (~1 octave) overall shift in spectral mean over the course of an attn-tono stimulus cycle that is unavoidable due to the constraints on the pairing of frequency bands.

Each of the twelve 9.6-minute-long runs was composed of eight 64s cycles plus 32-sec silent periods at the beginning and end of each run to allow for calculation of baseline auditory activation (Klein et al., 2014).

Behavioral Thresholds and Training. Participants first underwent behavioral tests of monaural pure-tone thresholds and binaural thresholds for detecting mini-sequences in quiet and in acoustic noise generated by the MRI scanner running the multiband EPI sequence. This provided a basis for adjusting center frequency amplitudes to approximate equal loudness in scanner noise. Participants also trained on the mini-sequence detection task in quiet and in acoustic scanner noise across two sessions.

Imaging Data Acquisition. Structural and functional images were acquired on a 3-Tesla Siemens Verio wide-bore MRI scanner at the Scientific Imaging and Brain Research (SIBR)

Center at Carnegie Mellon University using a phased array 32-channel head coil across three scan sessions on separate days. Stimulus presentation was under the control of a MacPro running PsychToolbox 3.0.12 in Matlab (The Mathworks, Inc.), with audio output to an external AD/DA converter (Babyface, RME) connected to an amplifier (Pylepro) that delivered stimuli to participants in the scanner diotically over MRI-compatible earbuds (Sensimetrics S14). All stimuli were pre-filtered to equalize sound stimuli according to the earbuds' frequency response. After participants were settled into the bore, sound volume was adjusted so that participants could comfortably hear all frequencies through scanner noise. Participants wore a fiber optic response glove (Current Designs) that communicated with a Brain Logics I/O device (Psychology Software Tools, Inc); participants used the glove to respond to mini-sequence repeats using the right index finger. During all functional scans, subjects closed their eyes to reduce the potential for stimulus-correlated eye movements.

In the initial scanning session (~50 min), we acquired multi-parameter mapping (MPM) images for quantitative myelin mapping and structural identification of primary auditory cortex on an individual basis while participants watched a film. Proton density-weighted (PDw), T1-weighted (T1w), and magnetization transfer (MTw) images were acquired using an in-house 3D FLASH pulse sequence (voxel size: 0.8 x 0.8 x 0.8 mm^3, matrix = 320 x 280 x 208, TR = 25.0 ms, bandwidth 488 Hz/px, excitation flip angle: 6^o (PDw/MTw) or 21^o (T1w), slab rotation 30^o). To accelerate this high resolution acquisition, a partial Fourier acquisition (6/8 coverage) was used in the inner phase-encoded direction (RL) and parallel imaging was used along the outer phase encoding direction (AP), reconstructed using the GRAPPA algorithm (acceleration factor 2, 18 integrated auto-calibration lines) as implemented on the scanner platform. Four gradient echoes were acquired for each contrast (TE=2.5, 4.74, 6.98, 9.22 ms) after each excitation pulse and averaged to improve SNR (Helms et al., 2009). Each FLASH acquisition lasted 9 minutes 45 seconds. Quantitative R_1 (=1/T1) maps were estimated from the PDw and T1w images

317 according to the model developed by Helms et al. (Helms et al., 2008) including a correction for
318 RF transmit field inhomogeneities (Lutti et al., 2010) and imperfect spoiling (Preibisch and
319 Deichmann, 2009). The transmit field map was calculated using a 3D EPI spin-echo
320 (SE)/stimulated echo (STE) method (Lutti et al., 2010; 2012); FOV = 256 x 192 x 192 mm,
321 matrix = 64 x 64 x 48, TE = 53.14 ms, TM = 47.60 ms, TR = 500 ms, bandwidth = 2298, nominal
322 α varying from 135^{o} to 65^{o} in steps of 5^{o}, acquisition time 6 minutes) and was corrected for off-
323 resonance effects using a standard B0 field map (double gradient echo FLASH, 3x3x2 mm
324 isotropic resolution, whole-brain coverage).
325
326 The final two scanning sessions acquired functional data for four runs each of the tonotopy,
327 'stepped' attn-tono, and 'randomized' attn-tono conditions. The runs were interleaved across
328 conditions and designed to assess phase-encoded functional influences of selective attention
329 across frequency ('stepped' attn-tono), the functional response to identical acoustics without
330 systematic phase encoded shifts of attention ('randomized' attn-tono), and functional responses
331 to single frequency bands identical to the attended bands in attn-tono, with phase-encoded
332 steps through frequency and no distractor frequency bands (tonotopy). Across all functional
333 runs, participants engaged in detecting repeats (1-back) of the four-tone mini-sequences. Run
334 order was counterbalanced according to condition and whether the cycle involved steps up or
335 down in frequency.
336
337 Functional images were acquired using a T2*-weighted echo-planar imaging (EPI) pulse
338 sequence (44 oblique axial slices, in plane resolution 3 mm x 3 mm, 3 mm slice thickness, no
339 gap, repetition time TR = 1000 ms, echo time TE = 41 ms, flip angle = 61°, matrix size = 64 x
340 64, field of view FOV = 192 mm). All EPI functional scans were performed using 4x multi-band
341 acceleration (Feinberg et al., 2010; Feinberg and Setsompop, 2013). There were 584 repetitions

342 acquired per run, with the first 8 images discarded to allow for longitudinal magnetization to
343 arrive at equilibrium. Runs were pseudo-randomly ordered across participants.
344
345 **Image Preprocessing.** *Cortical surface creation, and mapping of R_1 values.* Each subject's
346 cortical surface was reconstructed from a contrast-optimized synthetic FLASH volume, created
347 with mri_synthesize in Freesurfer from scaled and truncated versions of the T1 and proton-
348 density volumes; another MPRAGE-like synthetic image was created for use with the automated
349 Freesurfer Talairach procedure. Both volumes were conformed to 1mm isotropic resolution and
350 used in a customized reconstruction pipeline version. (In particular, the subject's PD volume
351 was used to deskull the synthetic FLASH image using a 'shrink-wrap' technique (Dale and
352 Sereno, 1993). After inspection for reconstruction quality, R_1 values were resampled from 50%
353 cortical depth fraction to the subject's surface, and also morphed to the unit icosahedron for
354 cross-subject curvature-aligned cortical surface based averaging (Fischl et al., 1999).
355
356 *EPI processing.* Each functional image from both sessions was aligned to a reference volume
357 from the middle of the first run using AFNI's 3dvolreg; registration and motion correction
358 goodness were hand-checked for each run. The reference volume was aligned to the subject's
359 cortical surface using boundary-based registration in Freesurfer (Greve and Fischl, 2009),
360 verified using manual blink comparison, and applied to the volume-aligned EPI data for
361 resampling. EPI data were analyzed in native space without any spatial smoothing using both
362 Fourier and general linear methods.
363
364 **Experimental Design and Statistical Analysis.** As noted above, the fMRI experiment used a
365 'discrete' version of a traditional phase-encoded design, such that both Fourier-based and
366 general linear model approaches could be used. Fourier analyses were carried out in csurf
367 (http://www.cogsci.ucsd.edu/~sereno/.tmp/dist/csurf) with individual and group analysis methods

368 employed as previously described (Sereno et al., 1995; Sereno and Huang, 2006; Hagler et al.,
369 2007). Functional activation amplitude was estimated as the Fourier amplitude of the periodic
370 BOLD signal (proportional to percent response) at the frequency of the stimulus cycle (8
371 repetitions per run). An F statistic was calculated by comparing that amplitude to the average
372 amplitude of other 'noise' frequencies (Hagler et al., 2007). Periodic signal components with
373 very low frequencies (due to slow head motion) and the second and third harmonic of the
374 stimulus were excluded as neither signal nor noise (this is mathematically equivalent to first
375 linearly regressing out these frequencies as nuisance variables before calculating significance).
376 The phase of the signal, which corresponds to a particular point of the stimulus cycle, was then
377 mapped to a color scale and the amplitude of the signal at each vertex was mapped to color
378 saturation (Gouraud sharing within each face). Runs with downward frequency steps were time-
379 reversed and averaged with upward-stepped scans in order to cancel fixed voxel-specific delays
380 in the BOLD response.
381
382 Linear modeling was carried out in FSL (Smith et al., 2004). For all runs, the motion-registered
383 data were high-pass-filtered (100 sec) and prewhitened; a hemodynamic model corresponding
384 to each stimulated and attended (tonotopy condition) or attended (stepped, randomized attn-
385 tono conditions) frequency band was created by convolving the 12.8-sec block with a gamma
386 function (lag 6s, SD sc). In a separate multiple regression, the unattended (ignored) frequency
387 band was modeled for both stepped and randomized attn-tono conditions. The verbal cue was
388 also modeled; all models were temporally filtered before multiple regression. Coefficients from
389 the first-level contrasts for each of the four runs were combined in a fixed-effects analysis for
390 each condition; data from the stepped and random block conditions were also combined in an
391 eight-run average.
392

Cross-subject averaging of phase-encoded mapping data was performed using the methodology developed by Hagler and Sereno (2006) in which the real and imaginary components of the signal with respect to the stimulus ramp are averaged across subjects, preserving any phase information consistent between subjects. This was performed by projecting each participant's phase-encoded map to the FreeSurfer spherical atlas using mri_surf2surf, performing 1 step of surface-based smoothing (< 1 mm FWHM in 2D), averaging across subjects at each vertex, then painting back onto a single subject's surface for viewing. For the multiple regression analyses, the same sampling process was used to sample each subject's contrast parameter estimates for cross-subject averaging and t-tests.

Surface-based cluster exclusion was used to correct for multiple comparisons in the group-wise averages (surfclust and randsurfclust from (Hagler et al., 2006)). The exclusion criterion (only surface clusters > 78 mm^2 unless otherwise noted) was determined based on the minimum estimated cortical area from iterative random sampling of cluster sizes (N=10000 iterations per hemisphere in randsurfclust) required to achieve a corrected alpha of $p < 0.001$ for each hemisphere, based on an initial uncorrected alpha of vertex-wise $p < 0.01$.

As an alternative means of defining primary auditory cortex, we projected the Morosan et al. (2001) 3D raw probability maps provided in the AFNI (version 16.3.13, Cox, 2012) to a FreeSurfer "fsaverage" brain registered to the Talairach target brain, resampled the data onto the cortical surface, and thresholded at $p > 0.30$ to create ROI labels. The labels were ~2 mm FWHM (five steps) surface-smoothed with manual removal of isolated marked vertices (due to "spillover" from the 3D to 2D projection within the lateral fissure), then spherically morphed to each subject. The labels were individually inspected (and filled if there were small holes in the label), then a boundary was delineated around each label on each subject's flattened auditory cortical patch.

ROI analyses. We quantified the similarity between frequency band response profiles driven by stimulus+attention (tonotopy) versus attention alone (attn-tono) in a 'quilt' of small cortical-surface-based ROIs that tiled the temporal plane. ROIs (as seen in Fig. 5b) were created on a single subject's right and left hemisphere flattened patches by flooding all vertices within a 4mm radius around a central selected vertex. Each of the ROIs (57 in the right hemisphere patch, and 68 in the slightly larger left patch) were then spherically morphed to the other 7 subjects' flattened patches. Spurious ROI sampling on the edges of the patches was manually corrected on the original subject's inflated surface and re-morphed to all other subjects. Each ROI was then projected into the registered native-space EPI volume using Freesurfer's mri_label2vol (sampled from the grey-white boundary to 0.8 of the calculated cortical depth, with fillthresh set to 0.5). For each subject, within each ROI, we calculated the average parameter estimate for each frequency band for tonotopy, and combined stepped and randomized attn-tono conditions. For each ROI, we then ran a linear model with average tonotopy parameter estimates for the 5 frequency bands predicting average attn-tono parameter estimates for the same bands, including subjects as a random factor. The resulting partial t-statistic for each ROI was z-transformed and color-rendered in Fig. 5b, with p-value thresholds Bonferroni-corrected to $p < 0.05$ for the number of ROIs per hemisphere, and indicated by the white outline surrounding the set of ROIs that surpass this z-threshold.

Results

Fourier-based Analyses

Stimulus- and attentionally-driven tonotopic organization in human auditory core. As a necessary first step, we characterized basic tonotopic (stimulus-driven) organization in and

immediately around myelin-estimated auditory core. The group-average R_1-based estimates of myelination (inflated hemispheres in left-most panel of Fig. 2) show that the highest R_1 values occur within primary somatomotor areas along the central sulcus, and in the typically keyhole-shaped presumptive 'auditory core' lying along and immediately surrounding Heschl's gyrus. It is important to note that myelination varies within auditory core and that the lateral and medial borders are less sharply demarcated (reviewed in Hackett et al., 2001, and Dick et al., 2012). To show this variation, we plot isointensity R_1 contours in the cortical flat patches in Figs 2a-c (with the curvature-based boundaries of Heschl's gyrus overlaid in dotted lines). To help identify the discontinuities in R_1 that would correspond to the putative borders of auditory core, we calculated the R_1 gradient along the surface (Glasser et al., 2016). Lines drawn along the peak gradient amplitude (not shown) corresponded well with the outermost R_1 isointensity contour in Fig. 2 (0.66 s^{-1}). (It is important to note the gradient at the lateral edge of presumptive core is quite shallow, and - as in postmortem myelin stains - it is therefore more difficult to establish an unambiguous lateral border, as could be surmised from the greater lateral spread of the isointensity contours. The shape and size (~1.2cm x 2.4cm) of presumptive auditory core in this sample also agreed with the results from Dick et al. (2012) at the same R_1 threshold (with the latter average core slightly narrower, not shown here).

The group-averaged topography of preferred frequency around auditory core has a typical arrangement (Dick et al., 2012; De Martino et al., 2015b), with the core surrounded by a high-frequency 'V'. Preferred frequency descends into the center of core (where R_1 values are highest) before reversing and slowly ascending to mid-frequency preferred frequencies anterolaterally (and to some extent posterolaterally). Fig. 3 shows tonotopic maps for each individual listener. In general, the relationship between auditory core and tonotopy group is conserved across listeners, but with some variability in the shape and extent of the isointensity R_1 contours. In particular, S2, S6 (right hemisphere), S7, and S8 (right hemisphere) had

irregularly shaped and 'blotchy' isointensity contours. While there is a fair degree of individual variability in results from human postmortem cyto- and myelo-architectonic studies of auditory core and surrounding areas (Hackett et al., 2001; Sweet et al., 2005), this was somewhat greater than expected variation given other work in our laboratory (Dick et al., 2012; Lutti et al., 2014; Carey et al., 2017). As an independent estimate of primary auditory areas, we also morphed the Morosan et al. (2001) 3D probabilistic map of primary auditory areas (TE1.0) using previously established methods (see Methods); the outlines of the morphed labels corresponding to $p > 0.30$ of being within TE1.0 are overlaid in black dotted lines in Figure 3.

We then asked whether 'attention-tonotopic' mapping resembled the tonotopic case in and around auditory core. Here, the group-level spatial distribution of tonotopy is closely recapitulated when spectrally-directed attention ('stepped' attn-tono condition) alone modulates activation (Fig. 2b). This holds true in and around the keyhole-shaped hyperintensity defining core, with a slight exception in the transition from higher to lower frequency preference in mid core. In contrast (and as expected) the group-level attn-tono response for the randomized control condition is much weaker (Fig. 2c), with almost no correspondence with the tonotopic map, despite being acoustically identical to 'stepped' attn-tono but for the shuffling of the verbal prompt ordering that destroyed the consistent phase-lag associated with specific frequency bands, and supporting these Fourier-based analyses. The one potential exception is in and around posterolateral core, where there is a low-to-mid frequency progression that is similar in attn-tono and tonotopic maps, particularly in the left hemisphere. (This may be due to the small (~1 octave) overall shift in spectral mean over the course of a stimulus cycle noted in Methods).

Stimulus- and attentionally-driven tonotopic organization outside of auditory core. In line with results from previous fMRI studies (Talavage et al., 2004; Woods et al., 2009; Humphries et al., 2010; Barton et al., 2012; Dick et al., 2012; Moerel et al., 2012; Saenz and Langers, 2014;

Thomas et al., 2015; De Martino et al., 2015b; Ahveninen et al., 2016; Leaver and Rauschecker, 2016; Riecke et al., 2016), there is stimulus-driven tonotopic mapping extending well beyond auditory core, spanning the temporal plane and continuing into the superior temporal sulcus (STS). As shown in Fig. 2a, the overall arrangement is characterized by two pairs of three interlacing best-frequency 'fingers,' with the high-frequency fingers (red/orange colormap) predominating medially and extending laterally, where they meet interdigitated lower-frequency fingers (green/yellow colormap) extending lateral to medial, with the longest 'middle' lower-frequency finger extending about halfway into auditory core. Similar to tonotopy within auditory core, the overall pattern of group activation can be observed in the majority of individual subjects (Fig. 3), but there is also considerable individual variability in the complexity, topography, and extent of tonotopic and attn-tono mapping, similar to that observed in the fMRI studies cited above (as well as electrophysiological studies in a number of studies in macaque and owl monkey, e.g., Merzenich and Brugge, 1973; Morel et al., 1993).

As can be seen in the maps in Fig. 2b, the tonotopically-aligned maps evoked by spectrally-directed attention are also present in the majority of auditory cortex outside of auditory core. Again, the structure of the tonotopic map (as revealed by Fourier analysis) is abolished when the attentional cue is randomized, thereby eliminating any consistent relationship between attended frequency band and phase lag (Fig. 2c).

The similarity between the maps evoked by presentation of a single frequency band (tonotopy) versus attention to one of two simultaneously-presented frequency bands ('stepped' attn-tono) can also be seen in each individual subject (Fig. 3). As with the group-averaged data, there is a close correspondence in the progression of preferred frequencies across auditory cortex in individual subjects. The similarity between the tonotopic and attn-tono maps is particularly striking in subjects 1, 2, 5, 6, and 7. The tonotopic organization of individual subjects

demonstrated overall commonalities, but with notable differences, even between individual subjects' right and left hemispheres, particularly outside of auditory core (Humphries et al., 2010; Moerel et al., 2014; Saenz and Langers, 2014). However, individual peculiarities were replicated across tonotopic and attn-tono conditions. In some subjects, there was a surprising lack of strong tonotopic mapping (subject 3 for which poor tonotopy may be due to greater EPI warping, and also subject 4 for which low frequencies dominated the tonotopic maps). In summary, there was a strong correspondence between tonotopic and attn-tono maps at both the group and individual levels.

Multiple regression analyses:

Winner-Takes-All: Maps of 'stepped' versus 'randomized' attention conditions, and quantitative concordance of tonotopic and attn-tono maps. In a complementary analysis, we used standard multiple regression techniques (see Methods) to estimate the BOLD response to each center-frequency band when it was presented in isolation (tonotopy), versus when it was attended in the presence of a distractor band (attn-tono). This allowed us to make use of the attentionally-driven signal in the randomized attn-tono condition and to combine these data with the results from the stepped attn-tono condition to increase statistical power. It also allowed us to verify that the attention effects generalize when listeners direct attention without the 'crutch' of consistent stepping up or down across attended frequency bands.

The auditory cortical patches in Fig. 4 show the cross-subject average 'Winner-Takes-All' (WTA) best frequency band (most positive-going BOLD response relative to resting baseline) maps for tonotopy and attn-tono conditions (with no shading for response amplitude). These are overlaid with the outermost R_1 isocontour (dashed yellow) corresponding to auditory core. As should be expected, the topography of the WTA maps essentially recapitulates the topography revealed

549 by the phase-encoded analyses. The same holds true of the attn-tono WTA maps from both the
550 'stepped' and importantly, the 'randomized' block conditions (Fig. 4); this result confirms that
551 even without the 'crutch' of the stepping frequency band, listeners can direct their attention to
552 specific frequency bands.
553
554 The WTA approach also allowed us to straightforwardly quantify the within-subject
555 correspondence between voxel-wise best frequency, as estimated by tonotopy and by attn-tono.
556 Here, we coded each voxel in native space as a '1' when best frequency was identical in both
557 conditions, and a '0' otherwise. We then resampled each subject's binary maps to their cortical
558 surface, and then averaged across subjects to create a 'concordance' map (Fig. 5a). These
559 maps (statistically thresholded at vertex-wise $p < 0.01$, with surface-cluster-corrected alpha of p
560 < 0.001) show that across subjects there was high concordance across best frequency maps
561 evoked by stimulus and by attention across much of the temporal plane in both hemispheres,
562 with little concordance in non-auditory areas. The extent of attentionally-driven tonotopic
563 mapping relative to overall tonotopicity is shown in the cortical patches below each concordance
564 map in Fig. 5a. Here, the outer contour of the significant ($p<0.001$ clusterwise corrected)
565 concordance map is overlaid on the phase-averaged group tonotopy map (same as Fig. 2a).
566 Averaging over all subjects, the majority of consistently tonotopically mapped cortex medial to
567 the crown of the superior temporal gyrus also shows preferred-frequency-aligned attn-tono
568 maps, as does a small posterior cluster. In the left hemisphere, almost all consistently mapped
569 tonotopic cortex also shows aligned attn-tono maps. However, it is important to note that there
570 are considerable individual differences in regional best-frequency alignment across the tonotopy
571 and attn-tono maps (as can be seen in Figure 3).
572
573 ***Comparison of response profiles to all frequency bands across tonotopy and attn-tono.***
574 As has been shown previously, e.g., Moerel et al., (2013), hemodynamic responses to

575 frequency in auditory cortex are not necessarily bandpass, but can be more complex and
576 multipeaked. Therefore, we also examined whether attention to a given frequency band in the
577 presence of a distractor band recapitulates the more graded response to non-preferred
578 frequencies observed when that frequency band is presented in isolation. To do this, we created
579 and surface-morphed a set of small cortical ROIs to each subject (see Fig. 5b and Methods),
580 and quantified the similarity between the tonotopy and attn-tono response profiles in each ROI
581 in each hemisphere by regressing the mean tonotopic parameter estimate for each frequency
582 band against the attn-tono parameter estimate (with subjects as a random factor). We used this
583 'ROI quilt' analysis (as opposed to a vertex-wise one) to capture regional variation in cross-
584 condition response profile similarity across subjects, which might be obscured by individual
585 differences in tonotopic map topography and surface-based registration errors. This also
586 reduced the number of statistical comparisons that must be corrected for, thus increasing power
587 to detect differences.
588
589 The region of interest (ROI) analyses (Fig. 5) further support the results from the 'concordance'
590 maps from the Winner-Takes-All analyses (Fig. 4). The ROI analyses (Fig. 5b) show that
591 individual subjects' tonotopy and attn-tono responses profiles are significantly associated across
592 most of auditory cortex (all ROIs within the white border), with the exception of the most lateral
593 aspects of the STG and upper bank of the STS. (Note that while there is a strong relationship
594 between tonotopy and attn-tono response profiles of each subject within a given ROI, there is
595 cross-subject variability in the particular shape of those response profiles, as suggested by the
596 individual maps in Fig. 3). There is a broad tendency for tonotopy/attn-tono profile similarity to
597 be strongest posterior-medially in both hemispheres, and no clear indication that profile
598 similarity is higher in auditory core (indeed, this is not the case in the left hemisphere). As
599 shown by the white line on the tonotopic flat maps in Fig. 5a, the area showing significant
600 response profile similarity extends over the majority of cortex showing strong tonotopic mapping

with topologic similarity across subjects. The response profile similarity extends into less tonotopically consistent regions medially and posteriorly, but does not include the more postero-lateral tonotopically mapped regions along the crown of the superior temporal gyrus.

Loser-Takes-All: Maps of 'dis-preferred' frequency. Given the graded nature of frequency response preferences we observed, we suspected that there would be a large-scale topography associated with the minimum BOLD response across frequency, and that this topography would also be recapitulated by attention. Thus, we also performed a parallel 'Loser-Takes-All' (LTA) analysis, in which we coded voxels by the frequency band driving the minimum BOLD response (again relative to resting baseline) and analyzed as above. The average descriptive LTA maps show roughly opposite frequency responses compared to the WTA tonotopic maps, with higher-frequency-band-preferring regions in the tonotopic map being least driven by lower-frequency-bands, and vice versa (Fig. 6a). There is also some overlap in the 'mid-frequency-preferring' regions, likely due to blurring of values when averaging subjects' integer-based maps. There is also quite close correspondence between the frequency band evoking the least response in the tonotopy (stimulus) condition and the smallest BOLD response evoked by attending to a given frequency band. The LTA concordance maps (Fig. 6b, statistical thresholding as in Fig. 5a) show that in the right hemisphere, the alignment of tonotopic and attn-tono maps is greatest in more lateral and anterior auditory cortex, with qualitatively somewhat greater concordance more medially in the left hemisphere. (Note that the hemispheric difference and also the apparent qualitative contrast with the WTA concordance maps seen in Figures 4 and 6 are exaggerated by the cluster-wise statistical thresholding combined with the overall slightly lower concordance in the LTA maps).

Difference in activation across auditory areas when best frequency is attended versus ignored. We also assessed the strength and consistency of BOLD-related frequency-band-

627 selective attention across subjects, and how the effect of attention varied with preferred or dis-
628 preferred frequency. We first used a subject's native-space WTA map to establish each voxel's
629 best frequency. Then, we assigned each voxel the parameter estimate for the difference in
630 activation between attending to its best frequency in the presence of a distractor, versus
631 attending to the distractor and ignoring its best frequency. (In other words, the value at each
632 voxel was the estimated difference in activation between attending to, versus ignoring, its best
633 frequency in the presence of other frequency bands). We repeated this process to estimate the
634 parallel attention effect for each voxel's 'worst' frequency (using the corresponding LTA map).
635 We then resampled each subject's native-space 'attention maps' to her/his cortical surface to
636 allow for surface-based cross-subject averaging and statistical testing (all again with a vertex-
637 wise $p< 0.01$ threshold and surface-cluster-corrected alpha of $p < 0.001$). The top row of Fig. 7
638 shows that across subjects there was significantly greater activation across most of auditory
639 cortex when best frequency was attended versus ignored. The widespread attention effect
640 included all of R_1-estimated auditory core (outlined in green), extending from the inferior circular
641 sulcus laterally to the upper bank of the STS, and antero-posteriorly from the temporal pole onto
642 the planum temporale. By contrast, there were relatively few regions where attention to a
643 voxel's least-preferred frequency band evoked greater activation than when the same
644 frequency band was the distractor. Attending to a voxel's least-preferred frequency
645 band only significantly increased activation along the posterior lateral STG in both
646 hemispheres, extending more medially in the left, and more anteriorly in the right (Fig. 7,
647 middle row). A direct comparison between these maps (cross-subject t-test on the difference
648 of differences, Fig. 7, bottom row) showed that there were considerable regional differences in
649 activation between attending to a voxel's preferred versus dis-preferred frequency band. In both
650 hemispheres, there was greater activation across most of the anterior temporal plane when
651 attention was directed to the preferred versus dis-preferred frequency; in the right hemisphere,

this effect extended throughout the temporal plane, as well as including a patch in the posterior STG. There were no regions in which the converse effect was observed (greater attend > distract activation for dis-preferred versus preferred frequency band). This shows that the frequency-selective attention-related BOLD gain is strongly modulated by frequency-preference - and provides some evidence for models of multiplicative, and not additive, attentional gain (but see Discussion).

Relationship of tonotopic and attn-tono map strength to MR-estimated myeloarchitecture. Typically, assays of cortical myelination are used to differentiate the most highly-myelinated cortical regions (like auditory core, MT/V5, or V1) from adjacent regions. This is true whether cortical myelination is assessed using ex-vivo 'gold-standard' approaches such as Gallyas staining, or estimated through *in-vivo* MRI T1-weighted/T2-weighted ratio, quantitative R_1, or magnetization transfer measures. However, more subtle myelination changes that occur throughout cortex may spatially correspond with changes in functional characteristics (Glasser et al., 2016; Wallace et al., 2016). For instance, recent combined fMRI and high-resolution quantitative MR show that slight reductions in cortical myelination in primary somatomotor cortex reliably occur at the border between face and hand areas (Kuehn et al., 2017).

Here, we asked whether the change in the degree to which cortex showed a strong frequency-band preference (namely, the amplitude of the phase-encoded tonotopic or attn-tono signal) spatially corresponded with changes in myelination as assessed by quantitative R_1 (within a 4 mm radius disk that roved across the entire cortical surface). The cross-subject-average normalized covariance map in Fig. 8a shows that there is a shared local gradient in tonotopic amplitude and R_1 along the entire inferior circular sulcus and the anterior part of the superior temporal gyrus, where tonotopic amplitude and R_1 drop in tandem over a narrow band of cortex.

678 There is also negative local spatial covariance between tonotopic amplitude and R_1 within the
679 center of auditory cortex, where tonotopic amplitude increases but R_1 remains relatively stable.
680 (There is also some tonotopic/R_1 spatial covariance within and around the central sulcus; these
681 regions showed considerably less overall amplitude in tonotopic response, but one that spatially
682 covaries with changes in R_1).
683
684 To test the replicability of this novel tonotopy-versus-R_1 searchlight cross-correlation, we
685 reanalyzed R_1 and tonotopy data from a previous study (Dick et al., 2012) that used a different
686 tonotopic stimulus (bandpass-filter-swept non-linguistic vocalizations) and a slightly different
687 multiparameter mapping protocol. Despite these methodological differences, we found a very
688 similar pattern of tonotopic/R_1 positive local spatial covariance within the circular sulcus and
689 along the lateral STG, with negative spatial covariance again in the center of auditory cortex
690 (Fig. 8b). The shared and relatively steep anterolateral and medial gradients in putative
691 myelination and degree of frequency specificity - observed in two independently-acquired
692 datasets - suggests a shared functional and myeloarchitectonic border, possibly similar in
693 character to those reported recently relating resting state, standard task activation, and T1-
694 weighted/T2-weighted derived myelination estimates across cortex (Glasser et al., 2016; Kuehn
695 et al., 2017).
696
697 As seen in Fig. 8c, the spatial relationship between local R_1 and *attn-tono* amplitude changes is
698 much less clear. Here, there is a weak relationship within and around auditory cortex that is only
699 observed within the circular sulcus (particularly in the right hemisphere). There are also stripes
700 of spatial covariation along the banks of the central sulcus, although not closely aligned with the
701 pattern observed with the tonotopy versus R_1 covariance maps. Although very preliminary,
702 these results suggest that changes in the degree of spectral attentional modulation in auditory
703 cortex are not strongly linked to the underlying myeloarchitecture, and stands in contrast to the

consistent spatial association in lateral and medial auditory cortex between local changes in R_1 and the strength of stimulus-driven frequency response preference.

Summary. Everyday listening ordinarily takes place in rich soundscapes within multiple, simultaneous sound sources contributing to the overlapping mix of sound waves that arrives at the ears. Auditory attention is crucial to sorting out the mix. Listeners direct attentional focus to a sound source, or even to specific acoustic dimensions within a single sound source, to zero in on auditory information that is diagnostic in guiding behavior.

We asked how endogenous attention directed to specific acoustic frequency bands modulates human auditory cortical activity. Using high-resolution quantitative MRI and a novel fMRI paradigm for driving sustained selective attention within specific frequency bands, we established effects of spectrally-specific attention in myeloarchitectonically-estimated human auditory core. These effects extend across the majority of tonotopically-mapped auditory cortex, and are apparent in individual listeners. Sensory-driven best-frequency tonotopic maps align with attentionally-driven maps across much of the temporal plane, with poor concordance in non-auditory areas. Individual tonotopic and attn-tono maps show correlated idiosyncracies. The frequency bands that evoke the least BOLD response from input and from attention also exhibit close spatial correspondence. There is greater activation across most of auditory cortex when best frequency is attended, versus ignored. Finally, there is local spatial correspondence in multiple auditory regions between the degree of R_1-estimated myelination, and the strength of the frequency-band-selective fMRI response for tonotopic stimuli.

Discussion

729 ***Human auditory core exhibits attentionally-driven tonotopic organization.*** Previous
730 findings showed similar stimulus- and attentionally-driven frequency preference in and around
731 Heschl's gyrus, a macroanatomical landmark associated with primary auditory areas (Da Costa
732 et al., 2013; Riecke et al., 2016). Here, we demonstrate that, within quantitative-R_1-defined
733 primary auditory areas, the attentionally-driven maps in each hemisphere are very similar to the
734 detailed tonotopic maps in the same subjects. As shown by comparison maps across the
735 acoustically-identical stepped and randomized attn-tono conditions (Figs. 2 and 4), the
736 alignment between tonotopic and attention maps depends on allocation of attention to the cued
737 frequency band, not perceptual interference or other stimulus-driven effects. The fact that there
738 is considerable, high-level attentional modulation within primary auditory areas is interesting
739 given previous results suggesting more limited attentional topographic modulation in primary
740 auditory (Atiani et al., 2014) and visual (Saygin and Sereno, 2008) cortex, compared to more
741 robust attentional modulation in areas immediately adjacent to primary ones.
742
743 ***Attentionally-driven tonotopic organization extends across much of auditory cortex.*** We
744 also find strong evidence for tonotopically-mapped spectrally-directed attention in much of
745 auditory cortex, particularly along the lateral superior temporal gyrus (potentially analogous to
746 lateral auditory belt and parabelt regions in macaque (Hackett, 2007)). In addition to the
747 concordance in and around auditory core, the most consistent group-level alignment of these
748 maps lies lateral to auditory core, with each map characterized by three higher-to-lower best-
749 frequency-band traversals, moving from posterior to anterior roughly along the superior
750 temporal gyrus.
751
752 This pattern suggests a cross-species parallel to results reported in ferret (Atiani et al., 2014),
753 where task-evoked attentional modulation of frequency-tuned neurons is particularly strong in
754 non-primary (dPEG) tonotopically mapped auditory areas in ferret. In this regard, the stimulus

755 complexity, variability, and memory demands of the current task may have helped to drive
756 attentional response in these more lateral and anterior areas. Our results are consistent with a
757 human fMRI comparison of cross-modal attentional effects (Petkov et al., 2004), which showed
758 greater activation in lateral auditory regions when attention was directed to a demanding
759 auditory repetition detection task than when the same sounds were played as subjects
760 performed a demanding visual detection task. However, our results differ from these studies to
761 some degree in that attentionally-driven tonotopic modulation in auditory core was also robust
762 (similar to cross-modal attention studies in macaque A1 (O'Connell et al., 2014) and primary
763 auditory areas (De Martino et al., 2015a)), and did not differ significantly from that in lateral belt.
764
765 There was good correspondence between the voxel-wise best frequency-band for tonotopy and
766 attention-tonotopy in individual listeners. Like several prior studies (data and review in
767 Humphries et al., 2010; Moerel et al., 2014; Saenz and Langers, 2014; Brewer and Barton,
768 2016; Leaver and Rauschecker, 2016), we observed quite substantial variation in the detailed
769 topography of tonotopy across individuals (but cf. Ahveninen et al., 2016). It is especially
770 noteworthy that attention-tonotopy recapitulated these topographic idiosyncrasies (as observed
771 in the concordance analyses, Fig. 4b and 5b).
772
773 It is intriguing that there was a systematic frequency-band-associated topography not only of
774 best frequency but also of dis-preferred frequency and, also, that the frequency-selective
775 attenuation of BOLD gain relative to other frequencies can be recapitulated by selective
776 attention to that frequency band in the presence of other spectral information. One could
777 speculate that this map structure might be a population-level reflection of an ‘inhibitory surround’
778 structure observed in some electrophysiology studies (Calford and Semple, 1995; Sutter et al.,
779 1999; but cf Wehr and Zador, 2003), with the frequency band driving the least BOLD response
780 corresponding to the deepest trough in an asymmetric surround -- an effect that could drive the

781 very similar tonotopic and attn-tono graded frequency response preferences revealed in the
782 multiple ROI analysis (Fig. 5b).
783
784 Here, the average frequency response profile evoked by the single-band tonotopic stimuli was
785 recapitulated by attention to the same frequency bands in the context of distractors. Prior
786 human neuroimaging research has been consistent with the possibility that the shape of the
787 frequency response in and around Heschl's gyrus is attentionally-modulated in a bandpass
788 manner that relies on amplification rather than attenuation (Riecke et al., 2016). Based on
789 results from a larger number of spectral bands, the current findings suggest that, at least at a
790 more macroscopic scale, spectrally-directed attention modulates cortical activity in a more
791 graded fashion, with the shape of the attentional response to both preferred and less-preferred
792 frequency bands similar to that evoked by stimulus alone - a contention supported by the
793 alignment of the 'Loser-Takes-All' tonotopic and attn-tono maps (Fig. 6). That is, the frequency
794 band that drives the smallest fMRI response when presented alone is also the frequency band
795 that elicits the least activation when attended in the presence of a distractor. A better
796 understanding of the mechanisms underlying these maps will require more fine-grained
797 characterization of frequency-directed attentional modulation, preferably at very high spectral
798 and temporal resolution (Moerel et al., 2013; Lutti et al., 2014; Moerel et al., 2014; Ahveninen et
799 al., 2016) that might also help to unveil cortical-depth-specific attentional effects (De Martino et
800 al., 2015a). In particular, it will be important to see whether different fMRI tasks - using more
801 complex naturalistic sounds, or more or less abstract cues to frequency - can mimic the task-,
802 valence- and context-dependent effects observed in non-human animal cortical auditory
803 receptive fields, where the character of the 'contrast-enhancing' modulations differs markedly
804 with experimental manipulation (Fritz et al., 2005; 2007a; 2007b; David et al., 2012; Atiani et al.,
805 2014; Kuchibhotla et al., 2017). (It is worth noting that task-related modulation of frequency-
806 selective attentional effects has long been of interest in human auditory psychophysics

807 (Greenberg, 1968; Scharf et al., 1987; Scharf, 1989; Moore et al., 1996; Green and McKeown,
808 2001)).
809
810 ***There is correspondence between local change in R_1-estimated myelination and the***
811 ***strength of fMRI-assessed relative frequency selectivity.*** We found that the change in the
812 degree to which a small (4 mm radius) patch of cortex shows strong frequency preferences in
813 tonotopy was positively spatially correlated with its degree of myelination as estimated by R_1.
814 The strength of the correlation was anatomically specific, marking the medial border of auditory
815 cortex (within the circular sulcus) and revealing a potential anatomical index of 'processing style'
816 (from more to less tonotopically mapped) along anterolateral superior temporal gyrus. We found
817 this pattern to hold true in the data from the current study as well as in an independent cohort
818 scanned with quite different tonotopic stimuli and with multiparameter maps acquired on a
819 different scanner model, with different sequence settings (Fig. 8c). Although there was a
820 relatively reliable pattern of R_1-tonotopy correspondence at a group level, there was some
821 notable individual variation in local shared R_1/tonotopy gradients relative to gyral anatomy.
822 Thus, these patterns may be more useful than curvature for establishing areal borders on an
823 individual subject basis, particularly when there is no obvious sharp change in a single measure
824 (for discussion see also Glasser et al., 2016). Such work holds promise for generating novel
825 hypotheses for more intensively characterized species like mouse, ferret, or marmoset,
826 particularly in tandem with imaging techniques that that can cover multiple cortical areas
827 simultaneously.
828
829 **Future Directions.** In the current study, we limited our investigation to broadly defined auditory
830 cortex, where there was good evidence for systematic tonotopic representation from a number
831 of previous studies (Talavage and Edmister, 2004; Hackett, 2007; Moerel et al., 2013; 2014;
832 Saenz and Langers, 2014; Leaver and Rauschecker, 2016). In future research it will be

informative to examine interactions with several frontal regions whose potential analogues are known to have direct feedforward and feedback connections in macaque monkeys (Romanski and Goldman-Rakic, 2002), and where in ferret there are clear modulatory influences on primary and non-primary auditory cortex during learning (Atiani et al., 2014; Shamma and Fritz, 2014). Similar to recent work in vision (Klein et al., 2014; Puckett and DeYoe, 2015), it will also be useful to establish the shape of the attentional population receptive field, and how this varies across auditory areas and relates to stimulus-driven auditory population receptive field size (Thomas et al., 2015). Finally, following on from our own pilot work, it will be exciting to explore whether higher-level auditory regionalization may follow along some of the 'fault lines' revealed by shared local tonotopic and myelin gradients, and whether or not more sophisticated and fine-grained spectral attentional manipulations may reveal a relationship between the degree of attentional malleability and underlying cortical architecture and circuitry.

References

Ahveninen J, Chang W-T, Huang S, Keil B, Kopco N, Rossi S, Bonmassar G, Witzel T, Polimeni JR (2016) Intracortical depth analyses of frequency-sensitive regions of human auditory cortex using 7TfMRI. NeuroImage 143:116–127.

Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, Fritz JB (2014) Emergent Selectivity for Task-Relevant Stimuli in Higher-Order Auditory Cortex. Neuron 82:486–499.

Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M (2016) Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc Natl Acad Sci USA 113:E616–E625.

Barton B, Venezia JH, Saberi K, Hickok G, Brewer AA (2012) Orthogonal acoustic dimensions define auditory field maps in human cortex. Proc Natl Acad Sci USA.

Brewer AA, Barton B (2016) Maps of the Auditory Cortex. Annu Rev Neurosci 39:385–407.

Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. J Neurophysiol 73:1876–1891.

Callaghan MF, Helms G, Lutti A, Mohammadi S, Weiskopf N (2014) A general linear relaxometry model of R1 using imaging data. Magn Reson Med 73:1309–1314.

Carey D, Caprini F, Allen M, Lutti A, Weiskopf N, Rees G, Callaghan MF, Dick F (2017) Quantitative MRI provides markers of intra-, inter-regional, and age-related differences in young adult cortical microstructure. bioRxiv, 139568.

868 Cox RW (2012) AFNI: what a long strange trip it's been. NeuroImage 62:743–747.

869 Da Costa S, van der Zwaag W, Marques JP, Frackowiak RSJ, Clarke S, Saenz M (2011)
870 Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. Journal of
871 Neuroscience 31:14067–14075.

872 Da Costa S, Van Der Zwaag W, Miller LM, Clarke S, Saenz M (2013) Tuning in to sound:
873 frequency-selective attentional filter in human primary auditory cortex. Journal of
874 Neuroscience 33:1858–1863.

875 Dale AM, Sereno MI (1993) Improved localization of cortical activity by combining EEG and
876 MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci
877 5:162–176.

878 David SV, Fritz JB, Shamma SA (2012) Task reward structure shapes rapid receptive field
879 plasticity in auditory cortex. Proc Natl Acad Sci USA 109:2144–2149.

880 De Martino F, Moerel M, Ugurbil K, Goebel R, Yacoub E, Formisano E (2015a) Frequency
881 preference and attention effects across cortical depths in the human primary auditory
882 cortex. Proc Natl Acad Sci USA 112:16036–16041.

883 De Martino F, Moerel M, Xu J, van de Moortele PF, Ugurbil K, Goebel R, Yacoub E, Formisano
884 E (2015b) High-Resolution Mapping of Myeloarchitecture In Vivo: Localization of
885 Auditory Areas in the Human Brain. Cerebral Cortex 25:3394–3405.

886 Dick F, Tierney AT, Lutti A, Josephs O, Sereno MI, Weiskopf N (2012) In Vivo Functional and
887 Myeloarchitectonic Mapping of Human Primary Auditory Areas. Journal of Neuroscience
888 32:16095–16105.

889 Dinse J, Härtwich N, Waehnert MD, Tardif CL, Schäfer A, Geyer S, Preim B, Turner R, Bazin PL
890 (2015) A cytoarchitecture-driven myelin model reveals area-specific signatures in human
891 primary and secondary areas using ultra-high resolution in-vivo brain MRI. NeuroImage
892 114:71–87.

893 Engel SA (2012) The development and use of phase-encoded functional MRI designs.
894 NeuroImage 62:1195–1200.

895 Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994)
896 fMRI of human visual cortex. Nature 369:525.

897 Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Glasser MF, Miller KL, Ugurbil K,
898 Yacoub E (2010) Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI
899 and Fast Diffusion Imaging Valdes-Sosa PA, ed. PLoS ONE 5:e15710.

900 Feinberg DA, Setsompop K (2013) Ultra-fast MRI of the human brain with simultaneous multi-
901 slice imaging. Journal of Magnetic Resonance 229:90–100.

902 Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a
903 coordinate system for the cortical surface. Hum Brain Mapp 8:272–284.

904 Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA (2010) Adaptive, behaviorally gated,
905 persistent encoding of task-relevant auditory information in ferret frontal cortex. Nature
906 Neuroscience 13:1011–1019.

907 Fritz JB, Elhilali M, David SV, Shamma SA (2007a) Does attention play a role in dynamic
908 receptive field adaptation to changing acoustic salience in A1? Hear Res 229:186–203.

909 Fritz JB, Elhilali M, Shamma SA (2005) Differential dynamic plasticity of A1 receptive fields
910 during multiple spectral tasks. J Neurosci 25:7623–7635.

911 Fritz JB, Elhilali M, Shamma SA (2007b) Adaptive changes in cortical receptive fields induced
912 by attention to complex sounds. J Neurophysiol 98:2337–2346.

913 Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson
914 J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC (2016) A multi-modal
915 parcellation of human cerebral cortex. Nature:1–11.

916 Green TJ, McKeown JD (2001) Capture of attention in selective frequency listening. J Exp
917 Psychol Hum Percept Perform 27:1197–1210.

918 Greenberg GZ (1968) Frequency-Response Characteristic of Auditory Observers Detecting
919 Signals of a Single Frequency in Noise: The Probe-Signal Method. J Acoust Soc Am
920 44:1513.

921 Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based
922 registration. NeuroImage 48:63–72.

923 Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in
924 auditory cortex of macaques, chimpanzees, and humans. J Comp Neurol 441:197–222.

925 Hackett TA (2007) Organization and Correspondence of the Auditory Cortex of Humans and
926 Nonhuman Primates. Evolution of the Nervous System, JH Kaas, Ed:109–119.Hagler D,
927 Saygin A, Sereno M (2006) Smoothing and cluster thresholding for cortical surface-
928 based group analysis of fMRI data. NeuroImage 33:1093–1103.

929 Hagler DJ Jr, Sereno MI (2006) Spatial maps in frontal and prefrontal cortex. NeuroImage
930 29:567–577.

931 Hagler DJ, Riecke L, Sereno MI (2007) Parietal and superior frontal visuospatial maps activated
932 by pointing and saccades. NeuroImage 35:1562–1577.

933 Helms G, Dathe H, Dechent P (2008) Quantitative FLASH MRI at 3T using a rational
934 approximation of the Ernst equation. Magn Reson Med 59:667–672.

935 Helms G, Draganski B, Frackowiak R, Ashburner J, Weiskopf N (2009) Improved segmentation
936 of deep brain grey matter structures using magnetization transfer (MT) parameter maps.
937 NeuroImage 47:194–198.

938 Herdener M, Esposito F, Scheffler K, Schneider P, Logothetis NK, Uludağ K, Kayser C (2013)
939 Spatial representations of temporal and spectral sound cues in human auditory cortex.
940 Cortex 49:2822–2833.

941 Herrmann B, Henry MJ (2013) Frequency-specific adaptation in human auditory cortex depends
942 on the spectral variance in the acoustic stimulation. Journal of ….

943 Humphries C, Liebenthal E, Binder JR (2010) Tonotopic organization of human auditory cortex.
944 NeuroImage 50:1202–1211.

945 Idemaru K, Holt LL (2011) Word recognition reflects dimension-based statistical learning. J Exp
946 Psychol Hum Percept Perform 37:1939–1956.

947 Kastner S, Ungerleider LG (2000) Mechanisms of visual attention in the human cortex. Annu
948 Rev Neurosci 23:315–341.

949 Klein BP, Harvey BM, Dumoulin SO (2014) Attraction of Position Preference by Spatial Attention
950 throughout Human Visual Cortex. Neuron 84:227–237.

951 Kuchibhotla KV, Gill JV, Lindsay GW, Papadoyannis ES, Field RE, Sten TAH, Miller KD,
952 Froemke RC (2017) Parallel processing by cortical inhibition enables context-dependent
953 behavior. Nature Neuroscience 20:62–71.

954 Kuehn E, Dinse J, Jakobsen E, Long X, Schäfer A, Bazin P-L, Villringer A, Sereno MI, Margulies
955 DS (2017) Body Topography Parcellates Human Sensory and Motor Cortex. Cerebral
956 Cortex:1–16.

957 Langers DRM, Krumbholz K, Bowtell RW, Hall DA (2014) Neuroimaging paradigms for tonotopic
958 mapping (I): the influence of sound stimulus type. NeuroImage 100:650–662.

959 Langers DRM, van Dijk P (2012) Mapping the tonotopic organization in human auditory cortex
960 with minimally salient acoustic stimulation. Cerebral Cortex 22:2024–2038.

961 Leaver AM, Rauschecker JP (2016) Functional Topography of Human Auditory Cortex. Journal
962 of Neuroscience 36:1416–1428.

963 Lutti A, Dick F, Sereno MI, Weiskopf N (2014) Using high-resolution quantitative mapping of R1
964 as an index of cortical myelination. NeuroImage 93 Pt 2:176–188.

965 Lutti A, Hutton C, Finsterbusch J, Helms G, Weiskopf N (2010) Optimization and validation of
966 methods for mapping of the radiofrequency transmit field at 3T. Magn Reson Med
967 64:229–238.

968 Lutti A, Stadler J, Josephs O, Windischberger C, Speck O, Bernarding J, Hutton C, Weiskopf N
969 (2012) Robust and fast whole brain mapping of the RF transmit field B1 at 7T. PLoS
970 ONE 7:e32379.

971 Merzenich MM, Brugge JF (1973) Representation of the cochlear partition of the superior
972 temporal plane of the macaque monkey. Brain Res 50:275–296.

973 Moerel M, De Martino F, Formisano E (2012) Processing of Natural Sounds in Human Auditory
974 Cortex: Tonotopy, Spectral Tuning, and Relation to Voice Sensitivity. Journal of
975 Neuroscience 32:14205–14216.

976 Moerel M, De Martino F, Formisano E (2014) An anatomical and functional topography of
977 human auditory cortical areas. Frontiers in Neuroscience 8.

978 Moerel M, De Martino F, Santoro R, Ugurbil K, Goebel R, Yacoub E, Formisano E (2013)
979 Processing of natural sounds: characterization of multipeak spectral tuning in human
980 auditory cortex. Journal of Neuroscience 33:11888–11898.

981 Moore BCJ, Hafter ER, Glasberg BR (1996) The probe-signal method and auditory-filter shape:
982 Results from normal- and hearing-impaired subjects. J Acoust Soc Am 99:542–552.

983 Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and
984 connections of auditory cortex in macaque monkeys. J Comp Neurol 335:437–459.

985 Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human
986 primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial
987 reference system. NeuroImage 13:684–701.

988 O'Connell MN, Barczak A, Schroeder CE, Lakatos P (2014) Layer Specific Sharpening of
989 Frequency Tuning by Selective Attention in Primary Auditory Cortex. Journal of
990 Neuroscience 34:16496–16508.

991 Oh J, Kwon JH, Yang PS, Jeong J (2013) Auditory imagery modulates frequency-specific areas
992 in the human auditory cortex. J Cogn Neurosci 25:175–187.

993 Paltoglou AE, Sumner CJ, Hall DA (2009) Examining the role of frequency specificity in the
994 enhancement and suppression of human cortical activity by auditory selective attention.
995 Hear Res 257:106–118.

996 Petkov CI, Kang X, Alho K, Bertrand O, Yund EW, Woods DL (2004) Attentional modulation of
997 human auditory cortex. Nature Neuroscience 7:658–663.

998 Pierpaoli C (2010) Quantitative Brain MRI. Topics in Magnetic Resonance Imaging, 21:63.

999 Preibisch C, Deichmann R (2009) Influence of RF spoiling on the stability and accuracy of T1
1000 mapping based on spoiled FLASH with varying flip angles. Magn Reson Med 61:125–
1001 135.

1002 Puckett AM, DeYoe EA (2015) The Attentional Field Revealed by Single-Voxel Modeling of fMRI
1003 Time Courses. Journal of Neuroscience 35:5030–5042.

1004 Rao AA, Talavage TM (2005) Reliability of phase-encode mapping in the presence of spatial
1005 non-stationarity of response latency. NeuroImage 28:563–578.

1006 Riecke L, Peters JC, Valente G, Kemper VG, Formisano E, Sorger B (2016) Frequency-
1007 Selective Attention in Auditory Scenes Recruits Frequency Representations Throughout
1008 Human Superior Temporal Cortex. Cerebral Cortex.

1009 Romanski LM, Goldman-Rakic PS (2002) An auditory domain in primate prefrontal cortex.
1010 Nature Neuroscience 5:15–16.

1011 Saenz M, Langers DRM (2014) Tonotopic mapping of human auditory cortex. Hear Res
1012 307:42–52.

1013 Saygin AP, Sereno MI (2008) Retinotopy and attention in human occipital, temporal, parietal,
1014 and frontal cortex. Cereb Cortex 18:2158–2168.

1015 Scharf B (1989) Spectral specificity in auditory detection: The effect of listening on hearing.
1016 Journal of the Acoustical Society of Japan (E).

1017 Scharf B, Quigley S, Aoki C, Peachey N, Reeves A (1987) Focused auditory attention and
1018 frequency selectivity. Percept Psychophys 42:215–223.

1019 Schönwiesner M, Zatorre RJ (2009) Spectro-temporal modulation transfer function of single
1020 voxels in the human auditory cortex measured with high-resolution fMRI. Proc Natl Acad
1021 Sci USA 106:14611–14616.

1022 Schwarzkopf DS, Song C, Rees G (2011) The surface area of human V1 predicts the subjective
1023 experience of object size. Nature Neuroscience 14:28–30.

1024 Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB
1025 (1995) Borders of multiple visual areas in humans revealed by functional magnetic
1026 resonance imaging. Science (New York, NY) 268:889–893.

1027 Sereno MI, Huang R-S (2006) A human parietal face area contains aligned head-centered
1028 visual and tactile maps. Nature Neuroscience 9:1337–1343.

1029 Sereno MI, Lutti A, Weiskopf N, Dick F (2013) Mapping the human cortical surface by combining
1030 quantitative T(1) with retinotopy. Cerebral Cortex 23:2261–2268.

1031 Shamma S, Fritz J (2014) Adaptive auditory computations. Curr Opin Neurobiol 25:164–168.

1032 Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H,
1033 Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J,
1034 Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and
1035 structural MR image analysis and implementation as FSL. NeuroImage 23:S208–S219.

1036 Stüber C, Morawski M, Schäfer A, Labadie C, Wähnert M, Leuze C, Streicher M, Barapatre N,
1037 Reimann K, Geyer S, Spemann D, Turner R (2014) Myelin and iron concentration in the
1038 human brain: A quantitative study of MRI contrast. NeuroImage.

1039 Sutter ML, Schreiner CE, McLean M, O'Connor KN, Loftus WC (1999) Organization of Inhibitory
1040 Frequency Receptive Fields in Cat Primary Auditory Cortex. J Neurophysiol 82:2358–
1041 2371.

1042 Sweet RA, Dorph-Petersen K-A, Lewis DA (2005) Mapping auditory core, lateral belt, and
1043 parabelt cortices in the human superior temporal gyrus. J Comp Neurol 491:270–289.

1044 Talavage TM, Edmister WB (2004) Nonlinearity of FMRI responses in human auditory cortex.
1045 Hum Brain Mapp 22:216–228.

1046 Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM (2004) Tonotopic
1047 organization in human auditory cortex revealed by progressions of frequency sensitivity.
1048 J Neurophysiol 91:1282–1296.

1049 Tardif CL, Gauthier CJ, Steele CJ, Bazin P-L, Schäfer A, Schaefer A, Turner R, Villringer A
1050 (2016) Advanced MRI techniques to improve our understanding of experience-induced
1051 neuroplasticity. NeuroImage 131:55–72.

1052 Tardif CL, Schäfer A, Waehnert M, Dinse J, Turner R, Bazin P-L (2015) Multi-contrast multi-
1053 scale surface registration for improved alignment of cortical areas. NeuroImage
1054 111:107–122.

1055 Thomas JM, Huber E, Stecker GC, Boynton GM, Saenz M, Fine I (2015) Population receptive
1056 field estimates of human auditory cortex. NeuroImage 105:428–439.

1057 Turner R (2015) Myelin imaging. In: Brain mapping: An encyclopedic reference, pp 137–142.
1058 Academic Press: Elsevier.

1059 Wallace M, Cronin M, Bowtell RW, Scott I, Palmer AR, Gowland P (2016) Histological Basis of
1060 Laminar MRI Patterns in High Resolution Images of Fixed Human Auditory Cortex.
1061 Frontiers in Neuroscience.

1062 Wehr M, Zador AM (2003) Balanced inhibition underlies tuning and sharpens spike timing in
1063 auditory cortex. Nature 426:442–446.

1064 Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X (2009)
1065 Functional maps of human auditory cortex: effects of acoustic features and attention.
1066 PLoS ONE 4:e5183.

1067

Figure Captions

Figure 1. Stimuli and design overview. (**A**) In a representative 12.8s **Tonotopy block** a neutral verbal prompt "*hear*" precedes 14 four-tone mini-sequences sampled around one of 5 center frequencies. The task is to detect the 1-3 mini-sequence repeats embedded within the block. The grey box highlights a single mini-sequence repeat. (**B**) A single **Attention-Tonotopy (Attn-tono) block** includes two simultaneous streams of mini-sequences with distinct center frequencies. Mini-sequence repeats occur in each stream. A verbal prompt ("*high*" or "*low*") directs listeners to attend to one stream and report mini-sequence repeats in that stream, while ignoring repeats in the unattended stream. Two randomly-ordered orientation tones at the center frequency of each stream alert listeners to the frequency neighborhood of the upcoming streams. (**C**) A single 64s **cycle of Stepped Attn-tono blocks** includes five 12.8s blocks that step up (shown), or down, in center frequency. In this single cycle**,** the frequency band to which attention is directed by the verbal prompt (indicated with "*high*"/"*low*" above each block) is acoustically matched with the Tonotopy cycle shown in (**B**), but there are always competing unattended mini-sequences in a distinct frequency band. (**D**) A single 64s **cycle of Randomized Attn-tono blocks** is acoustically identical to the Stepped Attn-tono cycle in (**C**), except that half of the verbal prompts have been swapped, and therefore no longer cue attention to frequency with consistent phase lags. (**E**) The distinction between Stepped and Randomized Attn-tono blocks is highlighted by examining the first three (of eight) **cycles of a Stepped (top) versus Randomized (bottom) Attn-tono run.** The focus of attention is color coded in the frequency-band-specific manner shown in (**A**). For the Stepped condition (top), there is a consistent relationship between the stimulus phase lag and the attended frequency across cycles within a run. Thus, for voxels that show a consistently higher response at one attended frequency band compared to all others, there will be a periodic response at 8 cycles/run at a given phase lag corresponding to the particular frequency band attended. For

the Randomized condition (bottom), there is no consistent relationship, providing a control condition for Fourier analyses because frequency-band-directed attention is aperiodic across a run. **(F)** The stimulus phase lag with the highest periodic BOLD signal amplitude is determined for each voxel, mapped to a color scale, and then painted onto the cortical surface patch. BOLD signal amplitude is mapped to the color's saturation. (Note: In these panels A-D, stimulus intensity is adjusted across the spectrum to aid visual presentation of energy across frequency bands. See Methods for details on actual intensity across frequency bands).

Figure 2. Group activation for Tonotopy and Attention-Tonotopy conditions, with R_1 contours showing putative auditory core. The left-most panel shows cortical-surface-based group-averaged R_1, projected on the lateral inflated surface of one subject. (The left hemisphere is mirror-reversed to align cortical maps for visual comparison). For tonotopic map display, a patch of cortex including the entire temporal plane (shown in purple on the inflated surface) was cut and flattened. Panels A-C show this region enlarged, with isocontour lines showing quantitative R_1 values for the group-averaged putative auditory core, and color maps showing group-averaged best frequency as a function of (**A**) Tonotopy, (**B**) Attn-tono (Stepped), and (**C**) Attn-tono (Randomized Control) conditions. The stars are fiduciary points to assist in visual comparisons of maps across conditions; the outline of Heschl's gyrus is in yellow dashed lines (in (**A**), from the individual subject whose cortical patch was used). Consistent with previous work, the tonotopic map is characterized by two pairs of three interlacing best-frequency 'fingers,' with the high-frequency fingers (red/orange colormap) showing greatest frequency preference medially and extending laterally, where they meet interdigitated lower-frequency fingers (green/yellow colormap) extending lateral to medial, with the longest 'middle' lower-frequency finger extending about halfway into auditory core. This pattern is evident in Fourier-analysis-derived maps of the 'stepped' Attn-tono condition but not in the 'randomized control' Attn-tono condition, for which the attentional response was phase-cancelled. All maps are

statistically masked by overall activation to tonotopy stimuli in each hemisphere (cluster-corrected $p < 10^{-8}$, and gently shaded to show relative amplitude).

Figure 3. Individual subjects' Tonotopy and Attention-Tonotopy maps. Each subject's Tonotopic and Attn-Tono (Stepped) Fourier-analysis-derived maps are displayed on the same subject's flattened superior temporal cortical patch. R_1 isocontours around presumptive auditory core are shown in white, with thick solid lines depicting the lowest-valued (outermost) R_1 isocontour, thin solid lines depicting the highest (innermost) R_1 isocontour, and dashed lines showing intermediate values. (R_1 values differ somewhat across individuals). Dashed black lines indicate the outline of the cortical-surface-morphed 'TE1.0' label, where the area inside the line contains vertices estimated to have a $p > 0.3$ probability of falling within primary auditory cortex based on the Morosan et al., (2001) postmortem probability atlas (see Methods). Activation maps are gently shaded to show changes in response amplitude, but are unthresholded for comparison with individual maps from previous studies, e.g., Da Costa et al., (2011).

Figure 4. Comparison of responses in regression-based 'Winner-Takes-All' maps, Tonotopy and Attn-Tono. The color maps projected onto the right *(top panels)* and left *(bottom panels)* hemisphere cortical patches (same as patches shown in purple in Figure 2) show the cross-subject average best frequency band ('Winner-Takes-All,' or WTA) for 'stepped' tonotopic (left) 'stepped' attn-tono (middle) and 'randomized' attn-tono conditions. Note that with the regression-based approach, the 'randomized' condition is also expected to evoke strong attentionally-driven tonotopic maps. The dotted yellow line depicts the outermost R_1 contour ($0.66\ sec^{-1}$) around presumptive auditory core as shown in Figure 2.

Figure 5. Comparison of Tonotopy and Attn-tono maps. (**A**) 'Concordance maps' are rendered in heatscale on the inflated hemispheres to illustrate the similarity in best frequency

between tonotopic and attn-tono maps (the latter averaged over stepped and randomized blocks). These maps were calculated in two stages. First, in each subject's native EPI space, a voxel was coded as '1' if tonotopy and attn-tono stimuli evoked the same best frequency, and otherwise coded as '0'. Second, for each subject, the concordance maps were resampled to the individual's cortical surface, and projected onto the unit icosahedron for cross-subject surface-based averaging, thereby creating a composite measure of agreement between tonotopy and attn-tono maps, weighted by the consistency of this agreement across subjects. The concordance maps are statistically masked with a cross-subject t-map, calculated versus chance agreement ($p=0.20$) with a surface cluster correction of $p < 0.001$ (vertex-wise $p<0.01$, cluster threshold surface area> 74 mm^2, (Hagler et al., 2006). To demonstrate the extent of tonotopically-mapped cortex that is similarly mapped through spectrally-directed attention, the phase-encoded tonotopy cortical patches from Figure 2a are overlaid with the outline of the thresholded concordance map shown by the yellow dotted line. The white solid outline shows the Bonferroni-corrected ROI-wise correspondence outline from the 'ROI quilt' in panel B **(B)** The shading in each small ROI patch shows the z-score for the partial fit between tonotopy and attn-tono responses to each frequency band (with subjects as a random factor). ROIs with significant z-scores (Bonferroni-corrected p-value threshold of $p < 0.05$) are indicated by the thin white outline.

Figure 6. Comparison of responses in regression-based 'Loser-Takes-All' maps, Tonotopy and Attn-tono. (**A**) The colormaps projected onto the same cortical patches as Figures 2 and 4 show cross-subject group-average maps that depict the frequency band that drives the *least activation* compared to all other frequency bands ('Loser-Takes-All', or LTA) in tonotopy and attn-tono (stepped plus randomized blocks) conditions and in right and left hemispheres. As in Figure 4, the presumptive auditory core shown by the dashed yellow line depicting the outermost R_1 contour (0.66 sec^{-1}). (**B**) The tonotopy versus attn-tono LTA

1172 concordance map was created as in Figure 5a; note that the midpoint of the heatscale has been
1173 lowered slightly compared to Figure 5a, reflecting the overall somewhat lower concordance in
1174 the LTA maps compared to WTA. The dotted yellow R_1 isocontour is the same as Figure 4.
1175
1176 **Figure 7. Comparison of maps when best frequency is attended versus ignored.** The
1177 heatscale (t-values, thresholded as in Figure 5a) depicts the cross-subject cortical-surface-
1178 based average difference in activation when the subject-specific best frequency band of each
1179 voxel was attended versus ignored. The dotted green R_1 isocontour estimating auditory core is
1180 as in Figure 4.
1181
1182 **Figure 8. Local normalized covariance between R_1 values and tonotopic and attn- tono**
1183 **response amplitude.** The heatscale value at each vertex represents the normalized spatial
1184 covariance within a 4mm (2D) radius between R_1 and the amplitude of the tonotopic or attn-
1185 tono signal (e.g., the amplitude of the Fourier component at the stimulus frequency of 8
1186 cycles/run). (**A**) The cross-subject (N=8) cortical-surface-based average normalized covariance
1187 between R_1 and tonotopic amplitude. (**B**) The R_1 versus tonotopy normalized covariance in an
1188 independent cohort (N=6), using data previously acquired with a different tonotopy protocol
1189 (bandpass-filter-swept non-linguistic vocalizations) and on a different scanner (Siemens 3T
1190 Trio); full protocol as described in Dick et al., (2012). (**C**) The average normalized covariance
1191 between R_1 and attn-tono amplitude in the current cohort.
1192