Author Manuscript

Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but dos not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: Population receptive field estimates of human auditory cortex.

Authors: Thomas JM, Huber E, Stecker GC, Boynton GM, Saenz M, Fine I

Journal: NeuroImage

Year: 2015 Jan 15

Volume: 105

Pages: 428-39

DOI: 10.1016/j.neuroimage.2014.10.060

Published in final edited form as:
Neuroimage. 2015 January 15; 105: 428–439. doi:10.1016/j.neuroimage.2014.10.060.

Population receptive field estimates of human auditory cortex

Jessica M. Thomas[1], **Elizabeth Huber**[1], **G. Christopher Stecker**[2,3], **Geoffrey M. Boynton**[1], **Melissa Saenz**[4,5], and **Ione Fine**[1]

[1] Department of Psychology, University of Washington, Seattle WA , USA 98195-1525

[2] Department of Speech and Hearing Sciences, University of Washington, Seattle WA, USA 98105

[3] Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville TN, USA 37232

[4] Laboratoire de Recherche en Neuroimagerie (LREN), Department of Clinical Neurosciences, Lausanne University Hospital, Switzerland 1011

[5] Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland 1015

Abstract

Here we describe a method for measuring tonotopic maps and estimating bandwidth for voxels in human primary auditory cortex (PAC) using a modification of the population Receptive Field (pRF) model, developed for retinotopic mapping in visual cortex by Dumoulin and Wandell (2008). The pRF method reliably estimates tonotopic maps in the presence of acoustic scanner noise, and has two advantages over phase-encoding techniques. First, the stimulus design is flexible and need not be a frequency progression, thereby reducing biases due to habituation, expectation, and estimation artifacts, as well as reducing the effects of spatio-temporal BOLD nonlinearities. Second, the pRF method can provide estimates of bandwidth as a function of frequency. We find that bandwidth estimates are narrower for voxels within the PAC than in surrounding auditory responsive regions (non-PAC).

Keywords

Auditory cortex; functional magnetic resonance imaging (fMRI); population receptive field (pRF); Tonotopy

Corresponding author: Jessica Thomas University of Washington, Department of Psychology, Guthrie Hall, Seattle WA 98195-1525 Box 351525; phone no. (206)-543-4095, jmthomas@uw.edu.

1 Introduction

Primate studies suggest that the “core” of the auditory cortex (AC), identified on the basis of the underlying cellular architecture, contains up to three subdivisions with borders delineated by tonotopic gradient reversals: A1, R, and RT (Hackett, 2008; Hackett et al., 1998). In humans, cytoarchitectonic and chemoarchitectonic studies have localized the AC core to approximately the medial two-thirds of Heschl's gyrus (HG) (Dick et al., 2012; Rademacher et al., 2001). Tonotopic organization within these areas has been measured using intracortical auditory event-related potentials (Liegeois-Chauvel et al., 1991), surgically implanted microelectrodes (Howard et al., 1996) and non-invasive functional imaging (Da Costa et al., 2011; 2013; Formisano et al., 2003; Humphries et al., 2010; Saenz and Langers, 2013; Striem-Amit et al., 2011; Talavage et al., 2004).

A number of functional MRI studies have identified two tonotopic gradients, thought to correspond to human homologues of areas A1 (hA1) and R (hR) that form mirror-image representations reversing at a low frequency border (Da Costa et al., 2011; 2013; Formisano et al., 2003; Humphries et al., 2010; Langers and van Dijk, 2011; Moerel et al., 2012; Striem-Amit et al., 2011). The low frequency region is found close to Heschl's gyrus, and some uncertainty remains as to the exact orientation of these two maps with reference to the gyrus (see Saenz and Langers, 2013 for a review).

Several tonotopy studies have relied on stimuli comprised of orderly ascending or descending frequency progressions (e.g. Da Costa et al., 2011; 2013; Striem-Amit et al., 2011; Talavage et al., 2004), and a phase-encoding analysis that identifies the “best frequency” of a voxel by either finding the phase of the sinusoid or time-lagged function that best-fits the voxel time course (Engel et al., 1994). While such methods provide a robust method for identifying tonotopically organized areas, several features of this approach contribute to uncertainty in interpreting results.

Frequency progression stimuli have the potential to cause habituation and/or expectation effects. Moreover, the “traveling wave” of BOLD activity induced across the cortical surface (Engel et al., 1994) by frequency progressions is likely modulated by spatiotemporal nonlinearities. Previous studies have shown that while the spatial and temporal summation of BOLD signals can be well approximated by the linear model; significant nonlinear spatiotemporal interactions do occur (Binda et al., 2013; Pihlaja et al., 2008; Zenger-Landolt and Heeger, 2003). Furthermore, estimates of best frequency based on a small number of presented frequencies tend be biased, especially near the edge of the stimulus range (Dumoulin and Wandell, 2008, Appendix B). Our data and that of Dumoulin et al. (Dumoulin and Wandell, 2008, Appendix B) suggest that these effects can be somewhat reduced by fitting data using a continuous function such as a Gaussian or a sinusoid capable of assigning voxels to outside the presented frequency range rather than using a “winner-take-all” method that assigns a value based on the highest correlation to the presented frequencies only, where voxels with a best frequency outside of the stimulus range are assigned to either the highest or lowest presented frequency. Thus, frequency progression stimuli, especially when analyzed using ‘winner take all’ methods, have the potential to result in an overrepresentation of frequencies near the beginning or the end of the sweep.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Analogous concerns have been described for visual retinotopic mapping methods (Binda et al., 2013; Dumoulin and Wandell, 2008; Duncan and Boynton, 2003; Haak et al., 2012), and recently discussed for tonotopic mapping methods (Langers et al., 2014a; 2014b).

Here, we estimate auditory frequency responses using a population Receptive Field (pRF) method developed by Dumoulin and Wandell (2008) that is less susceptible to many of these issues when utilizing a randomized stimulus, and has the added advantage of providing a means of estimating the receptive field size or bandwidth of individual voxels as a function of frequency.

2 Materials and Methods

2.1 Subjects

Four right-handed subjects (2 male, 2 female, ages 24-45) participated in two fMRI sessions. Subjects reported normal hearing and no history of neurological or psychiatric illness. All procedures, including recruitment, consenting, and testing, followed the guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the Institutional Review Board.

2.2 Stimulus Presentation

Auditory stimuli were generated in MATLAB using the Psychophysics Toolbox (www.psychtoolbox.org). Stimuli were delivered via MRI compatible insert earphones (S14, Sensimetrics), at a sampling rate of 44.1 kHz, with intensities adjusted to ensure flat frequency transmission from 100 Hz to 8 kHz. Subjects were instructed to keep their eyes closed during all scans.

We measured fMRI responses to two types of stimulation sequences: *ascending/descending tone progressions* and *random tone sequences*. Both stimuli were comprised of blocks of pure tone stimuli originally developed by Da Costa et al. (2011). Each frequency block lasted 2 s and contained eight pure tone bursts of the same frequency, with each burst lasting either 50 ms or 200 ms in duration (inter-stimulus interval = 50ms). Tone durations were alternated in pseudo-randomized order, switching durations at least 4 times during each 2 s block, resulting in a “Morse code” like pattern of long and short tones which served to increase the perceptual salience of the stimuli over the regular pattern of background scanner noise **(Figure 1A)**. The choice of 2 s for our frequency block duration was primarily motivated to facilitate comparison with a traditional progression paradigm used in previous studies. **Inline Supplementary Figure 1** shows the results of simulations that suggest when the effects of BOLD adaptation (Soltysik et al., 2004) are included, the optimal block duration would be between 1 to 5 s.

After sound system calibration, all stimulus sound intensities were adjusted according to a standard equal-loudness curve created for insert earphones (ISO 226) to approximate equal perceived loudness across all frequencies. Actual sound intensities (65-83 dB SPL) matched the perceived loudness of a 1000 Hz tone (reference frequency) at 70 dB SPL. Acoustic noise from the scanner was attenuated by expanding-foam eartips as well as protective ear muffs placed over the ear following earphone insertion. Subjects reported hearing all tones to be clearly and comfortably audible, and of roughly equal loudness across all frequencies.

2.2.1 Ascending/descending tone progressions—*Ascending/descending progressions* were identical to those of Da Costa et al. (2011). The frequencies used were: 88, 125, 177, 250, 354, 500, 707, 1000, 1414, 2000, 2828, 4000, 5657, and 8000 Hz (half-octave steps). Each frequency block was presented for 2 s before progressing to the next higher (*ascending progression*, **Figure 1B)** or lower (*descending progression*, **Figure 1C)** frequency until all 14 frequencies had been presented. This 28 s frequency progression was followed by a 4 s silent pause and this 32 s cycle was repeated 15 times per scan, for total scan duration of 8 min. For each subject we carried out one *ascending/descending* stimulation session, during which subjects listened to 4 *ascending progressions* and 2 *descending progressions*.

2.2.2 Random tone sequences—For the *random tone sequences* **(Figure 1D)**, stimuli were arranged into 240 equally spaced frequency blocks (2 s) ranging from 88-8000 Hz. Each block was presented only once per scan and the frequency block order was randomly shuffled for each scan. Following every 60 frequency blocks was a 12 s silent pause which allowed the pRF algorithm to estimate the baseline fMRI response in response to scanner noise alone (Zuiderbaan et al., 2012) to improve the accuracy of bandwidth measurements. The entire scan therefore lasted a total of 8 min 48 s. Each subject participated in a single *random tone sequence* scanning session, consisting of 6 scans, each containing a different random sequence of the same tones. Our goal in presenting random tone sequences was to reduce spatial and temporal correlations between neighboring frequencies in the stimulus sequence, thereby reducing the influence of spatiotemporal nonlinearities on pRF estimates.

2.3 Magnetic resonance imaging

Functional magnetic resonance images were acquired with a 3T Phillips Achieva scanner (Philips, Eindhoven, The Netherlands) at the University of Washington Diagnostic Imaging Sciences Center (DISC) using an 8-channel head coil. Foam padding minimized head motion.

2.3.1 Acquisition Sequences—A common issue with tonotopic mapping protocols is the reliability of frequency measurements estimated in the presence of loud acoustic scanner noise that can interfere with or mask the hemodynamic responses to presented stimuli (Langers et al., 2005). While sparse scanning techniques limit the effects of acoustic noise, they require a marked increase in the amount of scanning time needed as compared to continuous acquisition (Hall et al., 1999; Humphries et al., 2010; though see Petkov et al., 2006).

In order to examine the reliability of the pRF estimates calculated in the presence of acoustic scanner noise, data were gathered using two acquisition sequences: a standard EPI sequence and an attenuated EPI sequence designed with Philips SofTone software (SofTone factor of 4.0) to generate quieter scanner noise. Acoustic scanner noise was recorded from inside the coil for both acquisition types (for both 27 and 35 slices, since the slice number can also influence acoustic noise) using an MR-compatible fiber-optic microphone (FOMRI-II, Optoacoustics) placed within the scanner bore and saved in WAV format using Matlab. The overall A-weighted sound pressure levels (dBA) were also collected for both acquisition types using a B&K sound level meter model 2250.

A frequency spectrum of the recorded scanner noise was obtained for each acquisition type by performing a discrete Fast Fourier Transform and calculating the resulting frequency component magnitudes. The standard EPI sequence generated an overall louder (122 dBA) auditory scanner noise peaking at approximately 1000 Hz; while the attenuated EPI sequence generated quieter (105 dBA) scanner noise peaking at approximately 350 Hz, see **Inline Supplementary Figure 2**.

Each session contained six scans. Three functional scans were collected using the standardEPI sequence (35 slices, TR/TE = 2000/25 ms, flip angle = 80°, EPI-factor = 51, no slice gap). After discarding the first 5 timeframes of each functional scan, *ascending/descending* session scans consisted of 240 acquisitions at an effective voxel size of 2.0 × 2.0 × 3.00 mm^3 (FOV= 264 × 264 × 105 mm^3, matrix size= 132 × 132 × 35). *Random tone sequence* scans consisted of 260 volumes at an effective voxel size of 2.33 × 2.33 × 2.33 mm^3 (FOV = 224 × 224 × 81.55 mm^3, matrix size= 113 × 113 × 35).

The other three scans were collected utilizing the Philips SofTone parameter, which reduces acoustic noise by controlling the shape of the gradient waveform independent of the amplitude. This change in the gradient waveform required either the number of slices to be reduced or the TR to be lengthened. We chose to maintain the 2 s TR and reduce the number of slices while increasing the voxel size to maintain roughly similar volume coverage (27 slices, TR/TE = 2000/25 ms, flip angle = 76°, EPI-factor = 39, no slice gap). Both *ascending/descending* session and *random tone sequence* scans consisted of 260 volumes at an effective voxel size of 2.75 × 2.75 × 3.00 mm^3 (FOV = 220 × 220 × 81 mm^3, matrix size= 80 × 80 × 27).

Voxel volume differed across *ascending/descending* standard EPI sequences (12 mm^3) and *random tone* standard EPI sequences (12.65 mm^3). This change was made to provide a compromise voxel size which had an intermediate in-plane resolution (2.33 × 2.33 mm^2) in between that of the standard (2.0 × 2.0 mm^2) and attenuated (2.75 × 2.75 mm^2) EPI sequences, and was also chosen to create isotropic voxels. However, by replicating all analyses described below using only attenuated EPI acquisition sequences, we confirmed that differences in voxel size were not responsible for any observed differences between *ascending/descending* and *random tone sequences*.

For the attenuated sequence we used a significantly larger voxel volume (22.69 mm^3), which likely increased signal to noise (by a factor of ~1.6) at the cost of lower spatial resolution (Triantafyllou et al., 2005). However this decrease in non-physiological signal to noise seemed not to have a dramatic effect on pRF estimate reliability, see **Supplementary Table 2**.

2.3.2 MR Pre-processing—For analysis, data were resampled into 1 × 1 × 1 mm^3 volumetric space. Standard pre-processing of fMRI data was carried out using BrainVoyager QX software (version 2.3.1 Brain Innovation B. V., Maastricht, The Netherlands), including 3D motion correction and high-pass filtering (cut-off: 2 cycles per scan). Functional data were aligned to the T1-weighted anatomical image acquired in the same session (MPRAGE, 1 × 1 × 1 mm^3). The anatomical images acquired in the two sessions were aligned to each other and to each subject's 3D Talairach-normalized anatomical dataset. The BrainVoyager QX automatic segmentation routine was used to reconstruct the cortical surface at the white-gray matter border (with hand-editing to minimize segmentation errors) and the resulting smooth 3D surface was partially inflated.

2.3.3 Voxel Selection—For each subject, large regions of interest (ROIs) were selected from the partially inflated left and right hemisphere cortical surface meshes using drawing tools within BrainVoyager QX. ROI borders were drawn generously to include all voxels within a contiguous region of auditory cortex between the lateral border on the crown of the superior temporal gyrus, the medial border within the fundus of the lateral sulcus, the posterior border of the supramarginal gyrus, and the anterior border of the most anterior portion of the temporal lobe. Surface ROIs for each subject are shown here in **Inline Supplementary Figure 3**. These surface ROIs were then mapped back into the brain volume and expanded to include voxels from −1 to 3 mm around the gray-white matter boundary to assure that inaccuracies in the definition of the gray-white matter boundary did not affect the selection of voxels for data analysis. Rather than projecting this surface ROI into the functional volumes, we chose to index the voxels in anatomical space (at the cost of increased computation time). Preprocessed time-course data for each 3D anatomical voxel within the volume ROI were then exported to Matlab for further analysis.

2.3.4 PAC ROI—A functionally defined PAC ROI corresponding to hA1 and hR defined on the basis of the tonotopic gradients and informed by the underlying anatomy (illustrated by the solid black lines in **Figure 2, Figure 3, Figure 5**, and **Figure 7)**, was generated to compare bandwidth data across regions. A contiguous patch of cortical surface containing two primary tonotopic gradients centered on low frequencies on HG was manually selected for each hemisphere of individual subjects (Da Costa et al., 2011; 2013). Anterior and posterior borders drawn along the outer high-frequency representations were primarily based on the tonotopic maps generated using the *random tone sequences* and pRF analysis. Quantitative comparisons of the size and similarity of PAC ROI identifications when using either the phase-encoding analysis or the pRF analysis are shown in **Inline Supplementary Figure 4**. ROI similarity was assessed using the Jaccard similarity coefficient (Jaccard, 1912) according to the equation:

$$J(A,B)=\frac{|A\cap B|}{|A\cup B|} \quad (1)$$

where the size of the intersection between voxels included in ROI A and ROI B is divided by the size of the union of voxels included in ROI A and ROI B.

Lateral and medial borders were conservatively drawn to include only the medial two-thirds of Heschl's gyrus (Rademacher et al., 2001). Auditory responsive regions outside of this ROI are referred to as non-PAC.

2.4 Analyses 2.4.1 Phase-encoding analyses

Phase-encoding analyses on the *ascending/descending tone progressions* were based on Da Costa et al. (2011). Using Brain Voyager, a hemodynamic time course was predicted in response to the first 2 s sound block of each stimulus cycle, based on that individuals estimated hemodynamic response function (HDR, see below for how this was estimated). It was assumed that there was no response to the remainder of the stimulus cycle. This cyclical model function was shifted successively in time in 2 s increments (corresponding to the TR) to generate 14 time-lagged functions. Linear correlation was applied (between all 14 time-lagged functions and the measured fMRI time course) on a per-voxel basis. Each voxel was then color coded according to the lag function resulting in the highest correlation value with its time course (“winner-take-all”). Lag values were then separately averaged for *ascending progression* scans (4 scans) and the *descending progression* scans (2 scans) for each voxel within the ROI and were exported to MATLAB (Mathworks, version 7.11). As in the pRF analysis, only voxels with a correlation threshold of 0.10 were considered “tonotopic” and included for further analysis.

2.4.2 pRF analyses—Our pRF model computes the population receptive field for individual voxels based on a linear temporal model of the fMRI BOLD response time course. All population receptive field (pRF) analyses were carried out using custom software written in MATLAB.

The pRF analysis begins with a definition of the stimulus time course *s(f,t)*, provided to the model as a matrix of binary values marking the presence or absence of auditory stimulation over frequency and time. To generate a hemodynamically blurred stimulus time course *r(f,t)*, the stimulus time course was convolved with the estimated hemodynamic response function modeled as a gamma function *h(t):*

$$h(t)=\frac{\left(\frac{t-\delta}{\tau}\right)^{(n-1)}e^{-\left(\frac{t-\delta}{\tau}\right)}}{\tau(n-1)!} \quad (2)$$

with the initial parameters n=3, tau (τ)= 1.5, and delay (δ)=1.8, based on Boynton et al. (1996). The hemodynamic response function (HDR) parameters tau and delay were subsequently optimized for individual subjects (as described below).

$$r(f,t)=s(f,t)*h(t) \quad (3)$$

We estimated the population response using a one-dimensional Gaussian function *g(f)*, defined over log frequency. The frequency center (f_0) corresponds to the best frequency, while the standard deviation (σ) was used to estimate bandwidth by transforming the values into octaves and then calculating the full width half maximum (FWHM) of the Gaussian function. To create the predicted time series, we calculated the linear sum of the overlap between the input stimulus after hemodynamic blurring with the Gaussian receptive field for each basis function, *g(f)*:

$$p\left(t\right)=\int r\left(f,t\right)g\left(f\right)df \quad (4)$$

Model fits for each voxel were obtained by finding values that maximized the correlation between the predicted and actual fMRI time-courses. The initial parameters for frequency centers (f_0) spanned the range of the stimulus from 88 to 8000 Hz, and initial standard deviation (σ) values ranged from 0.5 to 4. The best fitting parameters from this set were then used as initial parameters for a nonlinear search algorithm (Matlab's *fminsearch* function) which uses unconstrained nonlinear minimization to find the pRF model parameters f_0 and σ that maximize the correlation between the pRF predicted time-series and the BOLD data.

We then estimated each individual subject's auditory HDR by holding f_0 and σ fixed and finding the best fitting parameters for τ and δ. To limit computational time this estimate was only carried out on a subset of voxels (1 out of every 6) within the anatomically defined ROI, after checking that restricting our estimation to a subset of voxels did not have an appreciable effect on the estimated HDR. Median (across all voxels with a fitted correlation value above 0.25) τ and δ parameters were used to provide the estimate of that individual's HDR. We then iteratively fit the pRF parameters (f_0 and σ) for all voxels within the ROI, using the individually fitted HDR parameters. Individual HDR parameters are reported in **Inline Supplementary Table 1**.

The procedure described above had some minor modifications from the original implementation of Dumoulin and Wandell (2008). First, similar to Zuiderbaan et al. (2012) we jointly fit individual HDR and pRF parameters using a single dataset (rather than estimating the HDR using a separate dataset). Second, we convolved the hemodynamic response function before calculating overlap with the pRF rather than afterwards. This is mathematically identical but significantly reduces computation time, since once the HDR parameters have been estimated the convolution of the stimulus image with the HDR only has to be carried out once. Finally, we maximized the correlation between the predicted, *p(t)*, and the obtained time courses rather than minimizing root mean squared error. The amplitude of each Gaussian pRF was estimated after fitting, based on the linear regression between the predicted and obtained time courses.

After fitting, only voxels that met the following criteria were retained for further analyses: (1) the correlation between the observed fMRI response time-course and that predicted by the best-fitting pRF (our goodness-of-fit index) was higher than 0.10 and (2) the standard deviation (σ) of the best fitting pRF was between 0.01-2 in log frequency space, or .0332-6.64 in octaves (chosen based on the resolution and the range of the frequencies presented). These limits were chosen because at bandwidths below 0.0332 the voxel would respond to a single presented frequency, and with bandwidths above 6.64 the voxel would respond similarly to all frequencies presented.

Unlike the “winner-take-all” phase-encoded analysis, the pRF method is not limited to fitting frequency centers within the range of frequencies (88 and 8000 Hz) presented in the stimulus. However frequency center estimates outside this range are likely to be somewhat inaccurate. We included well-fit voxels with frequency center values beyond the imposed frequency cutoffs in **Figure 2** and **Figure 3** and the histograms in **Figure 4**, but simply categorized them as high-pass (HP) or low-pass (LP).

2.4.3 Comparison of different acquisition sequences—The consistency across the two acquisition sequences was calculated as a Pearson correlation coefficient (r_{xy}) across the frequency center and bandwidth estimates from three *random tone sequence* scans of both acquisition types. We additionally computed mean (across subjects) "non-attenuated" correlation coefficients ($r_{x'y'}$) for both frequency center and bandwidth values according to the equation:

$$r_{x'y'} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}} \quad (5)$$

where r_{xx} and r_{yy} are the correlation coefficients across single scans of the same acquisition type. Non-attenuated correlation provides a measurement of the correlation across two different scan types having factored out measurement error as estimated using scans of the same acquisition type (Muchinsky, 1996; Spearman, 1904).

2.4.4 Effect of frequency center distribution on bandwidth estimates—Since pRF bandwidth estimates are dependent on the distribution of frequency centers of the underlying receptive fields, we estimated the effects of systematic frequency gradients and hemodynamic spatial blurring on bandwidth values estimated from the *random tone sequence* scans. For each subject, we estimated the slope of the change in frequency center within each voxel, based on the frequency centers of neighboring voxels. For every voxel for which the pRF model was successfully fit, we found the neighboring voxels within a 3mm radius (results were not highly dependent on the choice of radius). We then assumed that the frequency sweep within the voxel of interest consisted of a linear gradient between its best fitting frequency and those of its neighbors to predict the increase in bandwidth size due these factors.

2.4.5 Estimation accuracy—*A* bootstrapping procedure (to remove variability due to selection of the subset of scans to be included in the estimate) was carried out whereby the standard error of the frequency center or bandwidth was repeatedly estimated for 2-6 of the *random tone sequence* scans, irrespective of acquisition type. Estimation variability was calculated in terms of both standard error of the mean (SEM) and the relative standard error (RSE) of the mean frequency center or mean bandwidth estimate according to the equation:

$$RSE = \frac{100 \times STD}{\left(\bar{x} \ \sqrt{n}\right)} \quad (6)$$

Where STD is the standard deviation of the frequency center or bandwidth estimate (corrected for biases due to being a small sample using Cochran's theorem), *n* is the number of scans included in the estimate and is either the mean frequency center or mean bandwidth estimate (Cochran, 1934).

2.4.6 Double Gaussian model—

Data were also fit with a double Gaussian center surround model (Zuiderbaan et al., 2012). The surround bandwidth was constrained (using a cost function) to be larger than the center bandwidth and to be less than 6.64 octaves. The amplitude of the surround was not constrained to be negative. The difference in fit between the single and double Gaussian model was assessed using a nested model F-test. The original single Gaussian model had 2 free parameters (frequency center and bandwidth, with the amplitude fixed at 1) while the double Gaussian model had 2 additional parameters describing the width and the amplitude of the "surround".

3 Results

3.1 Tonotopic maps as a function of both stimulus and analysis methods

Figure 2 shows tonotopic maps using the *ascending* and *descending tone progressions* on the left hemisphere cortical surface mesh for two typical subjects. Phase-encoded maps were generated using either an average of four scans of the *ascending tone progressions* (left panel) or an average of two scans of the *descending tone progressions* (middle panel). Maps using the pRF method for *ascending* and *descending tone progressions* (right panel), were generated from all six (equally weighted) functional scans. Separate maps for the right hemisphere of these subjects are shown in **Inline Supplementary Figure 5**.

Tonotopic maps generated using the *random tone sequences* and the pRF method from an average of six (equally weighted) functional scans are shown for both hemispheres of all four subjects in **Figure 3**. All maps are based on scans from both the standard and attenuated EPI acquisition protocols since, as described below, frequency center estimates were highly correlated across the two types of acquisition protocol. The frequency center values of individual voxels are color-coded along a gradient, with red corresponding to the lowest frequency value (88 Hz) through blue corresponding to the highest frequency tested (8000 Hz), as in **Figure 1**. Low-pass (LP) voxels are colored dark red, while high-pass (HP) voxels are colored dark blue. Across all figures/analyses the same correlation threshold of 0.10 was used.

Across both stimulus types and analysis methods, clear mirror-symmetric tonotopic gradients, corresponding to the PAC subdivisions hA1 and hR, are visible. These mirror-symmetric gradients reversed at a low frequency border, centered either on the crown of HG (Subject 1 and 3 in **Figure 3**) or within the sulcus intermedius (SI) between the duplications of HG (Subject 2 and 4 in **Figure 3**). This low frequency region was flanked both caudally (hA1) and rostrally (hR) by high frequency regions creating a "high-low-low-high" gradient reversal extending across HG. In certain cases, areas hA1 and hR appear to be oriented in a V-shape (Subject 2 and 4 in **Figure 3**) with the high frequency endpoints situated more medially than the low frequency center. The maps generated for *ascending* and *descending progressions* and analyzed with either phase-encoding or pRF analysis are very similar, as can be seen in **Figure 2** and **Table 1**. While the general position of the gradient reversal was very consistent across *ascending/descending* and the *random tone sequence* stimuli, there are noticeable differences in the distribution of frequencies between the maps even when both are analyzed using the pRF method (**Figure 2** and **3** and **Table 1**).

Figure 4 shows stacked (across subjects) probability histograms of frequency center *(f_0)* values estimated for both stimulus types and analysis methods, collapsed across acquisition protocols. The histograms for the *ascending* and *descending tone progression* stimuli show a strong double peak with very strong representations of higher and lower frequencies – this double peak was visible when these data were analyzed using either “winner-take-all” phase encoding or the pRF analysis . This double peak was also robust to the choice of threshold – remaining equally salient across thresholds ranging between 0.1 (**Figure 4**) through 0.35. This double peak was even more noticeable in single scan data. Because the ”winner-take-all” model did not allow for lag values greater than 14 or less than 1, noise within individual scans is not zero mean. That pushes f_0 away from the boundary values (towards the center) when averaging across multiple scans. In contrast, the histogram for the *random tone sequence* stimuli has a different shape, with a heavier representation of mid-range frequencies. In addition, as might be expected from the approximate range of human hearing (20 Hz-20 kHz) and primate physiology data (Cheung et al., 2001) a reasonable proportion of voxels were characterized as high pass (HP), while few voxels were characterized as low pass (LP).

3.2 pRF measurements of bandwidth

Bandwidth measurements from *random tone sequences* are shown in **Figure 5**. Bandwidth estimates are not reported for the *ascending/descending sequences* for reasons discussed below. Bandwidth maps were generated from an average of all six *random tone sequence* functional scans using both the standard and attenuated EPI acquisition protocols since, as described below, the two types of acquisition protocol produced similar bandwidth estimates. For bandwidths between 1-8 octaves, color coding of the bandwidth map is linear. Narrow bandwidths (< 1 octave) were color coded as red, and broad bandwidths (> 8 octaves) were coded as blue.

We did not find a significant correlation between bandwidth and frequency center values (r = 0.046). Nor did we find a clear bandwidth gradient running orthogonal to the tonotopic gradient for any individual subjects (using a gradient analysis similar to that used by Petkov et al. (2006), data not shown). However, as shown in **Figure 6**, we did find that bandwidth values were smaller in the PAC than in non-PAC regions. A 2-way ANOVA testing Area (PAC vs. Non-PAC) × Hemisphere (Left vs. Right) found a significant main effect of Area [F (1, 12) = 16.36 p = 0.002] on mean bandwidth value. There were no significant main effect of Hemisphere [F (1, 12) = 0.54, p < 0.476] and no significant interaction of Area vs Hemisphere [F (1, 12) = 0.07, p < 0.8017]. While bandwidths were overall narrower in PAC than in surrounding non-PAC areas, our findings do not support the use of bandwidths as the sole criteria for defining PAC borders since there was significant heterogeneity of bandwidths within our tonotopically defined PAC ROI.

We estimated the effects of systematic frequency gradients and hemodynamic spatial blurring on bandwidth size. These estimates predicted that shifts in frequency preference as a function of distance across the cortical surface is likely to have increased our bandwidth estimates by a median value of ~10%.

Finally, we found a positive correlation between activation levels (GLM t-value, **Inline Supplementary Figures 6** and **7**) and bandwidth estimates (*random tone sequences* only), as shown in **Inline Supplementary Figure 8**. Broader tuning results in a higher activation level across the duration of the sound condition, leading to stronger responses in the sound vs. silence contrast.

3.3 Comparison of different acquisition sequences

Tonotopic maps and bandwidth maps looked very similar across standard and attenuated EPI sequences. **Figure 7** shows an example from Subject 3, generated from the *random tone sequences* and fitted using the pRF analysis technique. The scatter plots of **Figure 8** directly compare frequency center and bandwidth values between the two acquisition sequences across all voxels that met our selection criteria for all four subjects. Neither the maps, nor the scatter plots show any obvious mis-estimation of frequency values near the peak of the scanner noise. Individual subject correlation coefficients between three scans of the attenuated EPI sequence and three scans of the standard EPI sequences are shown in **Table 2**. The mean (across subjects) non-attenuated correlation coefficient (which represents the correlation after parsing out measurement error) between standard and attenuated sequences was 0.813 (varying between 0.742-0.965 across subjects) for frequency centers and 0.726 (varying between 0.566-0.813 across subjects) for bandwidth estimates. Mean correlation coefficients between single scans of the same or different type of EPI sequence are shown in **Inline Supplementary Table 2**.

While frequency center and bandwidth estimates were similar across the two acquisition sequences, further testing, including presenting stimuli using a sparse sequence, will be necessary to determine the extent to which acoustic scanner noise affects pRF estimation.

3.4 Estimation accuracy

Figure 9 shows pRF estimation variability as a function of the number of scans included in the estimate, for both frequency center (f_0) and bandwidth. PRF estimation variability is presented in terms of standard error (SEM) of the mean frequency center (f_0) or mean bandwidth estimate (leftward y-axis) and the percentage standard error (% STE) of the mean frequency center (f_0) or mean bandwidth estimate (rightward y-axis).

Frequency estimates were remarkably reliable – with only 2 scans (less than 9 minutes each) the estimated standard error was less than 5% of the estimated frequency center (f_0), suggesting that a single scan is adequate to obtain a basic tonotopic map. Unsurprisingly, we find that bandwidth estimates were less accurate than frequency estimates: for our dataset a full session was required to obtain bandwidth estimates with standard error variability of less than 25%. Although we corrected for overlap in the sampled distributions using the Cochran correction, given the small number of scans on which these simulations were carried out, these estimates should be considered a rough guideline and may underestimate variance.

3.5 Double Gaussian model

The difference between the single and double Gaussian model was assessed using a nested model F test that examined the percentage of voxels that were fit significantly better using a double Gaussian model. In 2 of the four subjects less than 5% of voxels were better fit by the double Gaussian model – suggesting that the model did not provide any additional explanatory power. For Subjects 1 and 3, 14% and 17% of voxels respectively were significantly better fit with the double Gaussian model. For voxels that were significantly better fit by the double Gaussian model, Subject 1 had a median improvement of 33% (goodness-of- fit value of 0.24 for the single Gaussian model and 0.32 for the double Gaussian model). Subject 3 had a median improvement of 36% (goodness-of- fit value of 0.22 for the single Gaussian model and 0.30 for the double Gaussian model). For these better fit voxels, the median amplitude of the inhibitory surround compared to the center amplitude (fixed at 1) was −0.51 for Subject 1 and −0.50 for Subject 3 (note that the surround was not constrained to be negative).

4 Discussion

We show here that a modified version of Dumoulin and Wandell's (2008) pRF technique can be used to generate tonotopic and bandwidth maps of human auditory cortex. When using traditional *ascending/descending tone progressions* this method results in tonotopic maps that are very similar to those obtained using traditional phase-encoding analysis methods. However, unlike phase-encoding methods, the pRF method does not require an orderly sequence; instead, functional maps of auditory cortex may be derived from responses to a wide array of stimuli, including random or pseudo-random sequences. While maps generated using *random tone progressions* identified roughly similar tonotopic areas as those obtained using *ascending/descending tone progressions*, the exact boundaries of these areas, and the representation of frequencies within these areas was somewhat different.

One advantage of the *ascending/descending tone progression* stimuli is that the maps generated from these stimuli are very robust. This makes progression stimuli well suited for determining whether an area is tonotopic and/or providing rough estimates of area boundaries based on frequency reversals when scan time is limited. However, as previously described, *ascending/descending tone progressions* are susceptible to habituation and/or expectation effects (Kastner et al., 1999) as well as other types of spatiotemporal nonlinearities (Pihlaja et al., 2008; Zenger-Landolt and Heeger, 2003) due to the “traveling wave” of BOLD activity on the cortical surface. This “traveling wave” improves signal to noise: goodness-of-fit index values were higher for *ascending/descending tone progressions* than for *random tone sequences*. However these effects also result in an over-emphasis of frequencies near the beginning or the end of the sweep (Binda et al., 2013; Dumoulin and Wandell, 2008; Duncan and Boynton, 2003; Haak et al., 2012; Langers et al., 2014a; 2014b). We therefore do not report bandwidth estimates for the *ascending/descending sequences* as these effects can result in an overestimation of receptive field size, probably due to spatiotemporal blurring. Additionally the *ascending/descending sequences* had a half-octave separation between frequency blocks. This under-sampling made it impossible to accurately estimate narrower bandwidths (estimated bandwidths became infinitely small). In contrast, *random tone sequences* do not suffer from these biases and likely provide a more accurate method for measuring the finer-scale tonotopic and bandwidth organization.

4.1 The organization of tonotopic gradients and tuning bandwidth

Like previous studies, we identified two tonotopic gradients, thought to correspond to hA1 and hR (Da Costa et al., 2011; Da Costa et al. (2013); Formisano et al., 2003; Langers and van Dijk, 2011; Moerel et al., 2012). These “high-low-low-high” mirror-symmetric gradients followed the morphology of the cortex, with the more posterior gradient corresponding to hA1 and the more anterior gradient corresponding to hR (Da Costa et al., 2011; 2013; Humphries et al., 2010; Langers and van Dijk, 2011).

Ours is the first fMRI study to examine bandwidth using the traditional psychoacoustic approach of randomly presented pure tones. A variety of plausible organizations for bandwidth have been suggested in the literature, for example: (1) bandwidth varies systematically with best frequency (Cheung et al., 2001; Recanzone et al., 1999), (2) bandwidth gradients run orthogonal to tonotopic gradient (Kayser et al., 2007a; Petkov et al., 2006), and (3) tuning widths are narrower within the PAC than surrounding non-PAC regions (Moerel et al., 2012; Petkov et al., 2006; Seifritz et al., 2006).

While we did not find a significant correlation between bandwidth and frequency center values nor a clear bandwidth gradient, we did find that bandwidth values were significantly smaller within PAC than in surrounding non-PAC regions. This is consistent with evidence from animal electrophysiology (Rauschecker et al., 1995; Tian and Rauschecker, 1994) and neuroimaging (Kayser et al., 2007a; Petkov et al., 2006). Additionally, a recent human study by Moerel et al. (2012) found some indication that bandwidths were narrower along HG. Here we find that the area of narrow bandwidths extends beyond the gyrus itself. However, there was some heterogeneity of bandwidths within our tonotopically defined PAC, with regions of broad bandwidth consistently contained within the PAC borders. Therefore, unlike Moerel et al. (2012) we do not limit the PAC borders to the areas of narrowest tuning bandwidth.

One caveat is that bandwidth estimates are more variable than frequency estimates, even when based on a full session of data. We believe this is not a limitation of our particular methods, since the reliability of our frequency estimates compare favorably with previous studies using similar techniques (Da Costa et al., 2011; Humphries et al., 2010; Moerel et al., 2012; Petkov et al., 2006; Striem-Amit et al., 2011). Thus, it is likely that additional methodological advances might reveal further topographical structure in bandwidth organization.

4.2 What do auditory population receptive fields represent?

It is not straightforward to extrapolate single neuron tuning-width data from responses within human fMRI voxels, since voxel-level BOLD signals include the responses of hundreds of thousands of neurons collectively across cortical layers and neuronal types (10^4-10^5 neurons per cubic mm in cortex). As described in Dumoulin and Wandell (2008), pRF estimates depend on the average receptive field of the population of neurons that drive the BOLD response. This population may be a biased sample of the underlying population. For example, the neurons driving the response may depend on the properties of the stimulus being used – it is possible that different pRF estimates will be obtained using band pass noise or naturalistic stimuli (Moerel et al., 2012) than with our pure tone frequency stimulus.

Dumoulin and Wandell (2008) found that pRF estimates of visual receptive field size were systematically larger than single unit recording estimates, but were relatively closely matched to local field potential (LFP) estimates of receptive field size. Similarly, our bandwidth estimates were generally larger than those from single-unit recordings, which have been shown to range between one tenth of an octave to wider than five octaves in primates (Cheung et al., 2001; Recanzone et al., 1999), but were consistent with those measured using auditory field potentials (Kayser et al., 2007b).

4.3 The use of forward models to examine auditory processing

Similar to other "forward models", the pRF method models the response properties of individual voxels, an approach that is becoming increasingly popular, see Naselaris et al. (2011) for a review. One advantage of such forward models is that they allow for flexibility in the choice of stimulus sequence. As previously mentioned, Moerel et al. (2012) used a forward model to examine tonotopic organization using naturalistic stimuli and found that the regions that showed preferential responses to human voice and speech were tuned to the lower range of the frequency spectrum. This tuning was apparent using both natural sound and simple tone stimuli. In our case we took advantage of this stimulus flexibility to examine tonotopic mapping using a more traditional psychoacoustic approach, using stimuli that consisted of randomly ordered pure tones.

A second advantage of forward models is that they allow for the modeling of tuning properties of individual voxels. In our case, because we were specifically interested in frequency tuning and bandwidth, we used a very simple model, assuming a Gaussian in log space. In contrast, Moerel et al. (2012) used a two-stage model that allowed them to directly estimate the spectral tuning of the voxel. However, to quantify frequency and bandwidth tuning a Gaussian was fit to this spectral response profile– thereby reducing their model to a close equivalent of ours. More recently, this same group did away with fitting a Gaussian to characterize multi-peaked spectral tuning profiles, as opposed to single bandwidth value (Moerel et al., 2013).

It is also possible to create more complex pRF models than a simple Gaussian – an approach intermediate between the spectral tuning model and the simple Gaussian pRF model we focus on in this paper. The improved fits for the double Gaussian model found within a subset of voxels in 2 of our 4 subjects suggest that some additional power may be obtained by permitting more complex frequency response profiles than a simple Gaussian (Zuiderbaan et al., 2012). Other interesting future directions will include presenting multiple tones at once, thereby allowing measurement of response saturation (Kay et al., 2013) and the inclusion of alternative temporal response profiles including onset and offset responses (Harms and Melcher, 2003). One of the advantages of the pRF approach is that it provides an elegant framework within which to assess whether more complex models can provide additional insight into auditory BOLD responses.

The pRF method described here also provides a natural methodology to examine differences in tuning as a function of various factors known to affect auditory processing such as attention (Da Costa et al., 2013), auditory training (Jäncke et al., 2001; Menning et al., 2000), musical experience (Pantev and Herholz, 2011), or loss of vision early in life (Elbert et al., 2002).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Binda P, Thomas JM, Boynton GM, Fine I. Minimizing biases in estimating the reorganization of human visual areas with BOLD retinotopic mapping. Journal of Vision. 2013; 13

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. The Journal of Neuroscience. 1996; 16:4207–4221. [PubMed: 8753882]

Cheung SW, Bedenbaugh PH, Nagarajan SS, Schreiner CE. Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. J Neurophysiol. 2001; 85:1732–1749. [PubMed: 11287495]

Cochran WG. The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Mathematical Proceedings of the Cambridge Philosophical Society. 1934; 30:178–191.

Da Costa S, van der Zwaag W, Marques JP, Frackowiak RSJ, Clarke S, Saenz M. Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. The Journal of Neuroscience. 2011; 31:14067–14075. [PubMed: 21976491]

Da Costa S, van der Zwaag W, Miller LM, Clarke S, Saenz M. Tuning In to Sound: Frequency-Selective Attentional Filter in Human Primary Auditory Cortex. The Journal of Neuroscience. 2013; 33:1858–1863. [PubMed: 23365225]

Dick F, Tierney AT, Lutti A, Josephs O, Sereno MI, Weiskopf N. In vivo functional and myeloarchitectonic mapping of human primary auditory areas. J Neurosci. 2012; 32:16095–16105. [PubMed: 23152594]

Dumoulin SO, Wandell BA. Population receptive field estimates in human visual cortex. NeuroImage. 2008; 39:647–660. [PubMed: 17977024]

Duncan RO, Boynton GM. Cortical magnification within human primary visual cortex correlates with acuity thresholds. Neuron. 2003; 38:659–671. [PubMed: 12765616]

Elbert T, Sterr A, Rockstroh B, Pantev C, Muller MM, Taub E. Expansion of the tonotopic area in the auditory cortex of the blind. J Neurosci. 2002; 22:9941–9944. [PubMed: 12427851]

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN. fMRI of human visual cortex. Nature. 1994; 369:525. [PubMed: 8031403]

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron. 2003; 40:859–869. [PubMed: 14622588]

Haak KV, Cornelissen FW, Morland AB. Population receptive field dynamics in human visual cortex. PLoS ONE. 2012; 7:e37686. [PubMed: 22649551]

Hackett TA. Anatomical organization of the auditory cortex. J Am Acad Audiol. 2008; 19:774–779. [PubMed: 19358457]

Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol. 1998; 394:475–495. [PubMed: 9590556]

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp. 1999; 7:213–223. [PubMed: 10194620]

Harms MP, Melcher JR. Detection and quantification of a wide range of fMRI temporal responses using a physiologically-motivated basis set. Hum Brain Mapp. 2003; 20:168–183. [PubMed: 14601143]

Howard MA 3rd, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Res. 1996; 724:260–264. [PubMed: 8828578]

Humphries C, Liebenthal E, Binder JR. Tonotopic organization of human auditory cortex. NeuroImage. 2010; 50:1202–1211. [PubMed: 20096790]

Jaccard P. The Distribution of the Flora in the Alpine Zone. New Phytologist. 1912; 11:37–50.

Jäncke L, Gaab N, Wüstenberg T, Scheich H, Heinze HJ. Short-term functional plasticity in the human auditory cortex: an fMRI study. Cognitive Brain Research. 2001; 12:479–485. [PubMed: 11689309]

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron. 1999; 22:751–761. [PubMed: 10230795]

Kay KN, Winawer J, Mezer A, Wandell BA. Compressive spatial summation in human visual cortex. J Neurophysiol. 2013; 110:481–494. [PubMed: 23615546]

Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. J Neurosci. 2007a; 27:1824–1835. [PubMed: 17314280]

Kayser C, Petkov CI, Logothetis NK. Tuning to Sound Frequency in Auditory Field Potentials. J Neurophysiol. 2007b; 98:1806–1809. [PubMed: 17596418]

Langers DR, Van Dijk P, Backes WH. Interactions between hemodynamic responses to scanner acoustic noise and auditory stimuli in functional magnetic resonance imaging. Magn Reson Med. 2005; 53:49–60. [PubMed: 15690502]

Langers DRM, Krumbholz K, Bowtell RW, Hall DA. Neuroimaging paradigms for tonotopic mapping (I): The influence of sound stimulus type. NeuroImage. 2014a; 100:650–662. [PubMed: 25069046]

Langers DRM, Sanchez-Panchuelo RM, Francis ST, Krumbholz K, Hall DA. Neuroimaging paradigms for tonotopic mapping (II): The influence of acquisition protocol. NeuroImage. 2014b; 100:663–675. [PubMed: 25067814]

Langers DRM, van Dijk P. Mapping the Tonotopic Organization in Human Auditory Cortex with Minimally Salient Acoustic Stimulation. Cerebral Cortex. 2011

Liegeois-Chauvel C, Musolino A, Chauvel P. Localization of the primary auditory area in man. Brain. 1991; 114(Pt 1A):139–151. [PubMed: 1900211]

Menning H, Roberts LE, Pantev C. Plastic changes in the auditory cortex induced by intensive frequency discrimination training. NeuroReport. 2000; 11:817–822. [PubMed: 10757526]

Moerel M, De Martino F, Formisano E. Processing of Natural Sounds in Human Auditory Cortex: Tonotopy, Spectral Tuning, and Relation to Voice Sensitivity. The Journal of Neuroscience. 2012; 32:14205–14216. [PubMed: 23055490]

Moerel M, De Martino F, Santoro R, Ugurbil K, Goebel R, Yacoub E, Formisano E. Processing of natural sounds: characterization of multipeak spectral tuning in human auditory cortex. J Neurosci. 2013; 33:11888–11898. [PubMed: 23864678]

Muchinsky PM. The Correction for Attenuation. Educational and Psychological Measurement. 1996; 56:63–75.

Naselaris T, Kay KN, Nishimoto S, Gallant JL. Encoding and decoding in fMRI. Neuroimage. 2011; 56:400–410. [PubMed: 20691790]

Pantev C, Herholz SC. Plasticity of the human auditory cortex related to musical training. Neurosci Biobehav Rev. 2011; 35:2140–2154. [PubMed: 21763342]

Petkov CI, Kayser C, Augath M, Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS Biol. 2006; 4:e215. [PubMed: 16774452]

Pihlaja M, Henriksson L, James AC, Vanni S. Quantitative multifocal fMRI shows active suppression in human V1. Hum Brain Mapp. 2008; 29:1001–1014. [PubMed: 18381768]

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K. Probabilistic mapping and volume measurement of human primary auditory cortex. NeuroImage. 2001; 13:669–683. [PubMed: 11305896]

Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. Science. 1995; 268:111–114. [PubMed: 7701330]

Recanzone GH, Schreiner CE, Sutter ML, Beitel RE, Merzenich MM. Functional organization of spectral receptive fields in the primary auditory cortex of the owl monkey. J Comp Neurol. 1999; 415:460–481. [PubMed: 10570456]

Saenz M, Langers DR. Tonotopic mapping of human auditory cortex. Hear Res. 2013

Seifritz E, Di Salle F, Esposito F, Herdener M, Neuhoff JG, Scheffler K. Enhancing BOLD response in the auditory system by neurophysiologically tuned fMRI sequence. NeuroImage. 2006; 29:1013–1022. [PubMed: 16253522]

Soltysik DA, Peck KK, White KD, Crosson B, Briggs RW. Comparison of hemodynamic response nonlinearity across primary cortical areas. NeuroImage. 2004; 22:1117–1127. [PubMed: 15219583]

Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904; 15:72–101.

Striem-Amit E, Hertz U, Amedi A. Extensive cochleotopic mapping of human auditory cortical fields obtained with phase-encoding FMRI. PLoS ONE. 2011; 6:e17832. [PubMed: 21448274]

Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM. Tonotopic Organization in Human Auditory Cortex Revealed by Progressions of Frequency Sensitivity. J Neurophysiol. 2004; 91:1282–1296. [PubMed: 14614108]

Tian B, Rauschecker JP. Processing of frequency-modulated sounds in the cat's anterior auditory field. J Neurophysiol. 1994; 71:1959–1975. [PubMed: 8064359]

Triantafyllou C, Hoge RD, Krueger G, Wiggins CJ, Potthast A, Wiggins GC, Wald LL. Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. Neuroimage. 2005; 26:243–250. [PubMed: 15862224]

Zenger-Landolt B, Heeger DJ. Response suppression in v1 agrees with psychophysics of surround masking. J Neurosci. 2003; 23:6884–6893. [PubMed: 12890783]

Zuiderbaan W, Harvey BM, Dumoulin SO. Modeling center-surround configurations in population receptive fields using fMRI. J Vis. 2012; 12:10. [PubMed: 22408041]

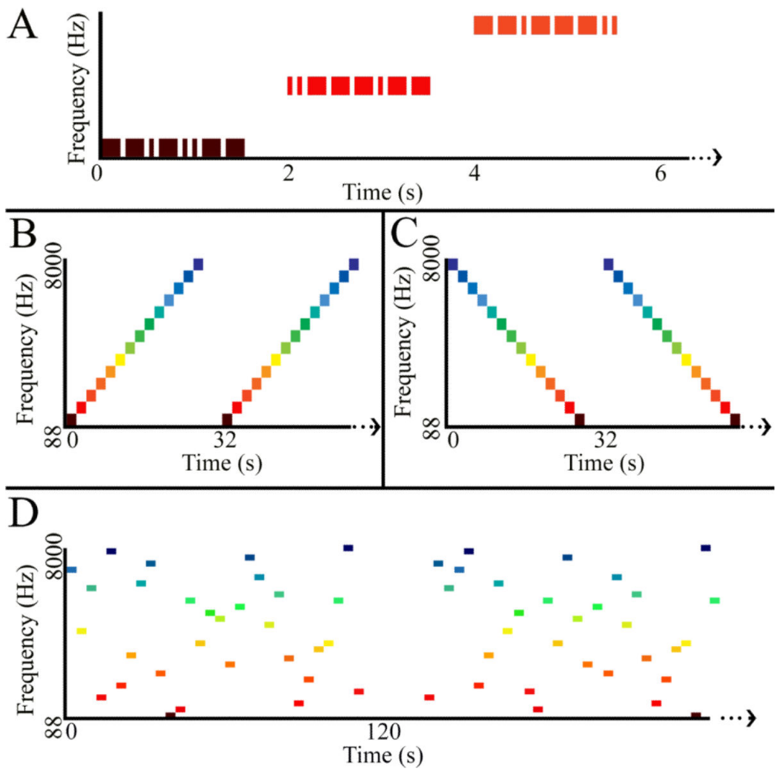

Figure 1.

Stimuli. (A) 3 example frequency blocks containing 8 pure tone bursts (50 or 200 ms). (B, C) 2 cycles (total 15 cycles per scan) of the *ascending and descending tone progressions*. (D) The first 120 frequency blocks of a *random sequence* scan.

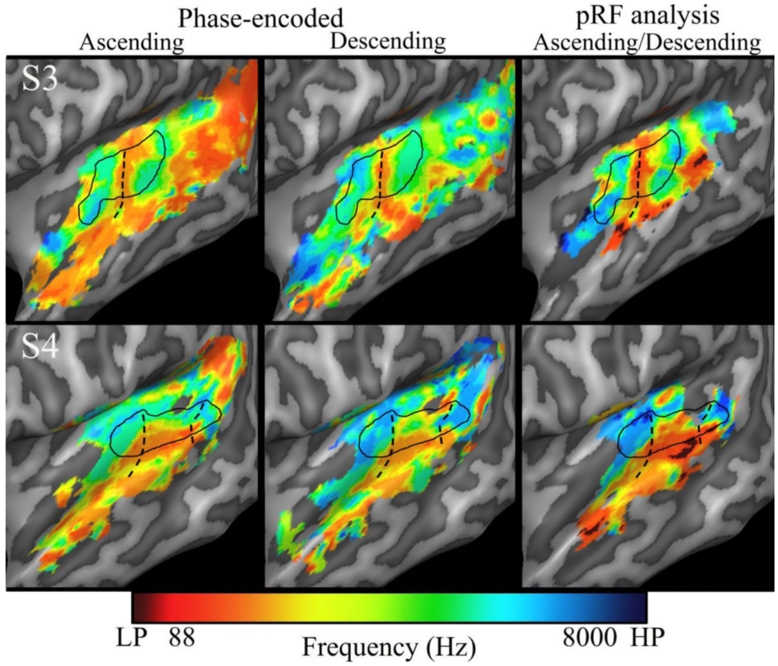

Figure 2.

Tonotopic maps for the left hemisphere of Subject 3 (first row) and Subject 4 (second row). Maps were calculated using either: *ascending tone progressions* (4 scans) analyzed using phase-encoding (left panels), *descending tone progressions* (2 scans) analyzed using phase-encoding (middle panels), or *ascending/descending tone progressions* (6 scans, therefore weighted towards the *ascending tone progressions*) analyzed using the pRF method (right panels). Frequency center (f_0) values are color-coded along a gradient, with red corresponding to the lowest frequency value (88 Hz) through blue corresponding to the highest frequency (8000 Hz). Low-pass (LP) voxels are colored dark red, while high-pass (HP) voxels are colored dark blue. The borders of the PAC ROIs are designated by the solid black lines. The crowns of gyri are indicated by the dashed black lines. Across all figures/ analyses the same correlation threshold of 0.10 was used, unless otherwise stated.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

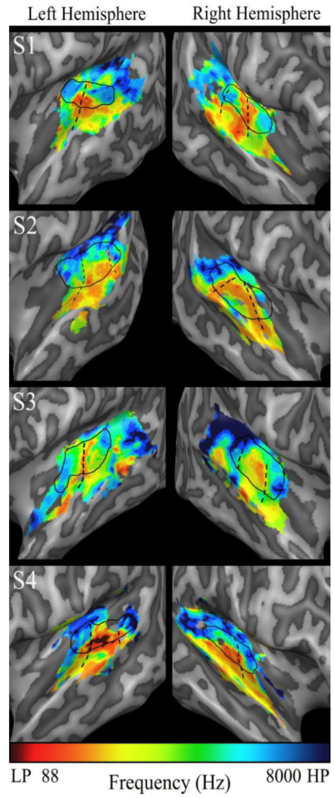

Figure 3.

Tonotopic maps calculated using *random tone progressions* analyzed using the pRF method. Frequency center (f_0) values are color-coded along a gradient, with red corresponding to the lowest frequency value (88 Hz) through blue corresponding to the highest frequency (8000 Hz). Low-pass (LP) voxels are colored dark red, while high-pass (HP) voxels are colored dark blue. The borders of the PAC ROIs are designated by the solid black lines. The crowns of gyri are indicated by the dashed black lines. Across all figures/analyses the same correlation threshold of 0.10 was used, unless otherwise stated.

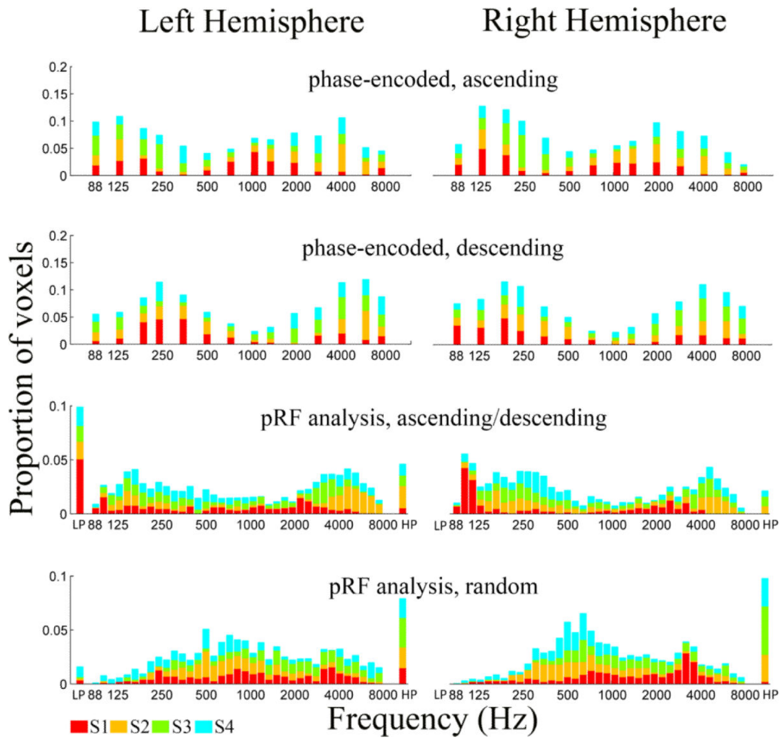

Figure 4.

Probability histograms of frequency center (f_0) values for voxels in left and right hemispheres, stacked across subjects for the tonotopic maps of **Figures 2** and **3** calculated for *ascending progressions* analyzed using phase-encoding (first panel), *descending progressions* analyzed using phase-encoding (second panel), *ascending/descending tone progressions* analyzed using the pRF method (third panel), or *random tone progressions* analyzed using the pRF method (fourth panel). The number of voxels is normalized so each subject contributes equally. The colors correspond to the individual subjects.

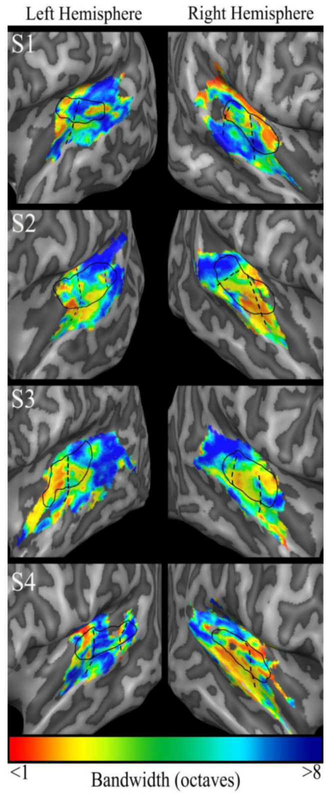

Figure 5.

Bandwidth maps for all four subjects. Bandwidth values are color-coded along a gradient with red corresponding to the narrow bandwidth values (<1 octave) through blue corresponding to the broader bandwidth values (>8 octaves).

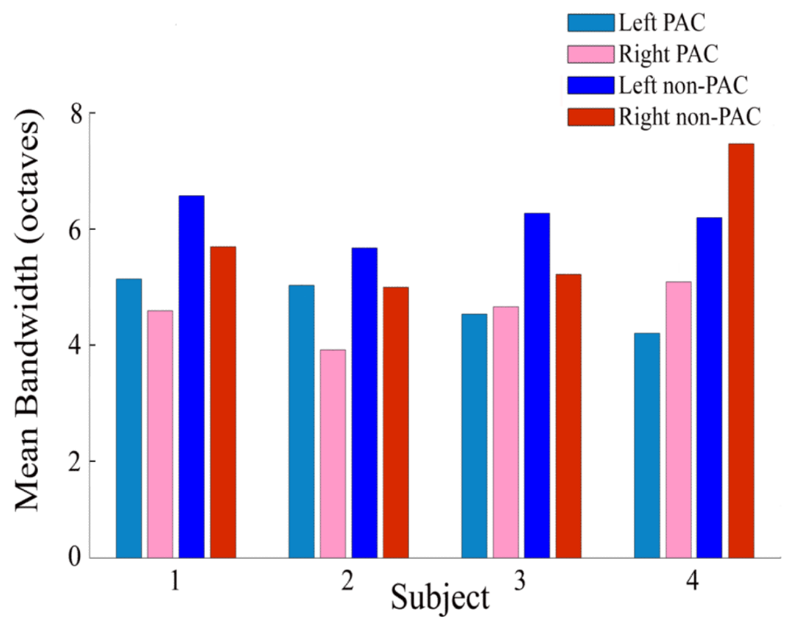

Figure 6.
Bandwidths are narrower within the PAC than in outlying non-PAC areas. As described in the main text, a 2-way ANOVA found a significant main effect of cortical area (PAC vs. non PAC) on mean bandwidth value.

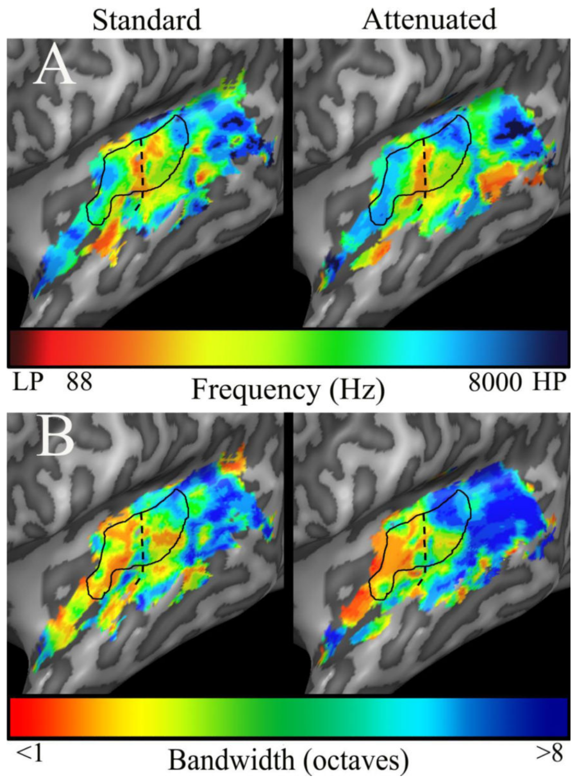

Figure 7.
(A) Tonotopic maps for example Subject 3 calculated using random tone progressions averaged across separate data sets of three scans using a standard EPI sequence and an attenuated EPI sequence, analyzed using the pRF method. Color coding is the same as **Figure 3.** **(B)** Bandwidth maps for for the same data and analysis. Color coding as in **Figure 5**.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

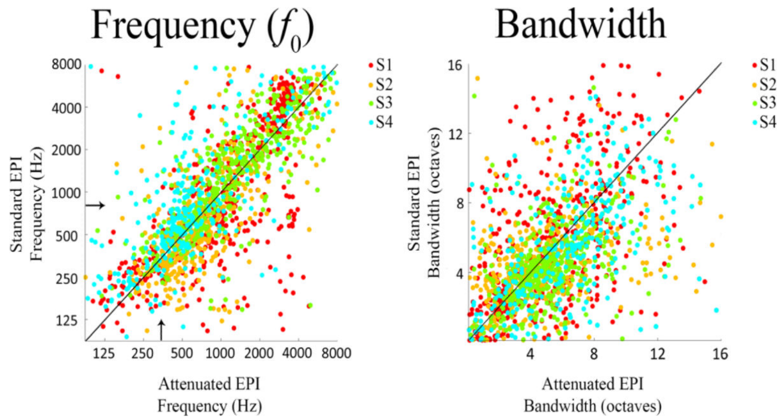

Figure 8.

Scatter plots showing the correlation of frequency center (f_0) (left panel) or bandwidth values (right panel) between the two acquisition sequences. Each data point represents a single voxel whose correlation (between the predicted and obtained time-course, calculated across the three scans of standard the attenuated acquisition protocols) was larger than our threshold of 0.1 for both acquisition sequences. The colors correspond to the individual subjects.

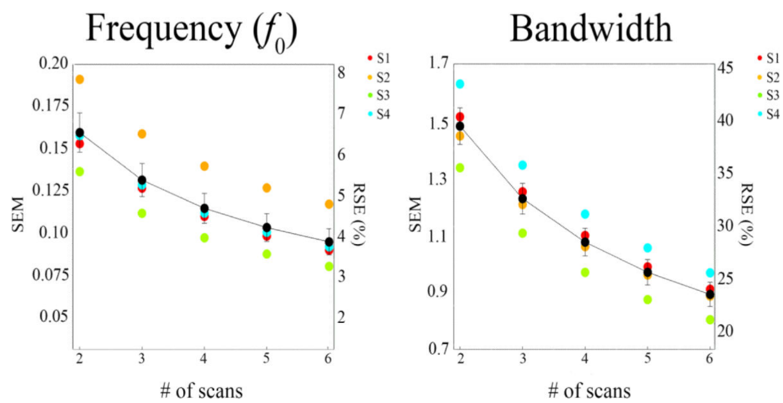

Figure 9.

Estimated standard error of frequency (f_0) (left panel) and bandwidth (right panel) estimates as a function of the number of scans. The x-axis represents the number of scans included in the estimate, the leftward y-axis represents the standard error of the mean, and the rightward y-axis represents the relative standard error, as described in the main text. The colored markers correspond to the individual subjects. The mean (across subjects) is shown in black and error bars represent standard error across subjects.

Table 1

Correlation coefficients between frequency center (f_0) values

Correlation coefficients between frequency center (f_0) values estimated using different stimulus sequences (A: *ascending progressions*, D: *descending progressions*, AD: *ascending/descending progressions*; R: *random tone sequences*) and analysis methods (PE: phase-encoding; pRF: pRF analysis). In each case correlation coefficients were calculated after having calculated frequency centers using scans collapsed across acquisition type. Separating data acquired using standard vs. attenuated acquisition EPI sequences did not change the pattern of results (data not shown).

	Subject			
	S1	**S2**	**S3**	**S4**
PE_A vs pRF_{AD}	0.713	0.857	0.682	0.861
PE_D vs pRF_{AD}	0.805	0.861	0.758	0.830
pRF_{AD} vs pRF_R	0.534	0.892	0.765	0.785

Table 2

Correlation coefficients for data collected using standard vs. attenuated acquisition protocols

Correlation coefficients for frequency center (f_0) values and bandwidth estimates from *random tone sequence* data collected using standard vs. attenuated acquisition protocols for all subjects. In each case correlation coefficients were calculated after having calculated pRFs using all three scans of a given acquisition type.

Subject	frequency (f_0)	bandwidth
S1	0.716	0.590
S2	0.745	0.401
S3	0.797	0.484
S4	0.711	0.560