

# Larger Auditory Cortical Area and Broader Frequency Tuning Underlie Absolute Pitch

 Larissa McKetton,<sup>1,2,4</sup> Kevin DeSimone,<sup>2,3,5</sup> and  Keith A. Schneider<sup>1,2,6</sup>

<sup>1</sup>Department of Biology, <sup>2</sup>Centre for Vision Research, <sup>3</sup>Department of Psychology, York University, Toronto, Ontario, Canada, M3J 1P3, <sup>4</sup>Joint Department of Medical Imaging, University Health Network, Toronto, Ontario, Canada, M5T 2S8, <sup>5</sup>Department of Psychology, New York University, New York, New York 10003, and <sup>6</sup>Department of Psychological and Brain Sciences, University of Delaware, Newark, Delaware 19716

Absolute pitch (AP), the ability of some musicians to precisely identify and name musical tones in isolation, is associated with a number of gross morphological changes in the brain, but the fundamental neural mechanisms underlying this ability have not been clear. We presented a series of logarithmic frequency sweeps to age- and sex-matched groups of musicians with or without AP and controls without musical training. We used fMRI and population receptive field (pRF) modeling to measure the responses in the auditory cortex in 61 human subjects. The tuning response of each fMRI voxel was characterized as Gaussian, with independent center frequency and bandwidth parameters. We identified three distinct tonotopic maps, corresponding to primary (A1), rostral (R), and rostral-temporal (RT) regions of auditory cortex. We initially hypothesized that AP abilities might manifest in sharper tuning in the auditory cortex. However, we observed that AP subjects had larger cortical area, with the increased area primarily devoted to broader frequency tuning. We observed anatomically that A1, R and RT were significantly larger in AP musicians than in non-AP musicians or control subjects, which did not differ significantly from each other. The increased cortical area in AP in areas A1 and R were primarily low frequency and broadly tuned, whereas the distribution of responses in area RT did not differ significantly. We conclude that AP abilities are associated with increased early auditory cortical area devoted to broad-frequency tuning and likely exploit increased ensemble encoding.

**Key words:** absolute pitch; auditory cortex; Heschl's gyrus; music; tonotopy; tuning sharpness

## Significance Statement

Absolute pitch (AP), the ability of some musicians to precisely identify and name musical tones in isolation, is associated with a number of gross morphological changes in the brain, but the fundamental neural mechanisms have not been clear. Our study shows that AP musicians have significantly larger volume in early auditory cortex than non-AP musicians and non-musician controls and that this increased volume is primarily devoted to broad-frequency tuning. We conclude that AP musicians are likely able to exploit increased ensemble representations to encode and identify frequency.

## Introduction

Absolute pitch (AP), also referred to as perfect pitch, is the ability to identify or recreate a given note or collection of notes in the absence of a reference note (Deutsch, 2013). It is not simply a better ability to hear, but the ability to mentally classify sounds into remembered categories. The prevalence of AP is relatively rare, with estimates of <1 in 10,000 persons reported (Bachem, 1955; Profita and Bidder, 1988; Deutsch, 2013), affecting both genders equally (Deutsch et al., 2006). Additionally, AP ability is

rare even among expert musicians who have had the same amount of musical training and have spent tens of thousands of hours practicing and reading scores (Deutsch et al., 2009). Many noted musicians, such as Mozart, Bach, and Beethoven, had AP, whereas many other equally prominent musicians, such as Wagner and Schumann, lacked it (Sacks, 2007).

The contributions of genetics and experience to the development of AP are still debated, but it seems that there is a critical period during which musical training must occur (Levitin and Zatorre, 2003; Russo et al., 2003; Miyazaki and Ogawa, 2006). Several brain imaging studies investigating AP have identified differences in cortical thickness and connectivity in sound-, music-, and memory-processing regions (Loui et al., 2011; Dohn et al., 2015). Increased activation has been observed in the left superior temporal sulcus in AP musicians during a pitch memory task compared with controls (Schulze et al., 2009) and increased functional activations in the STG, bilateral Heschl's gyrus (HG),

Received June 15, 2018; revised Jan. 8, 2019; accepted Jan. 12, 2019.

Author contributions: L.M. wrote the first draft of the paper; L.M. and K.A.S. edited the paper; L.M., K.D., and K.A.S. designed research; L.M. performed research; K.D. contributed unpublished reagents/analytic tools; L.M., K.D., and K.A.S. analyzed data; L.M., K.D., and K.A.S. wrote the paper.

The authors declare no competing financial interests.

Correspondence should be addressed to Keith A. Schneider at keithas@udel.edu.

<https://doi.org/10.1523/JNEUROSCI.1532-18.2019>

Copyright © 2019 the authors

and middle temporal gyrus (MTG) in the AP group compared with the control group during a music-listening task (Loui et al., 2012). However, the fundamental neural mechanisms underlying AP are not well understood.

We wondered whether differences in the precision of low-level frequency representation might account for the special abilities in AP. The human auditory cortex is organized into tonotopic maps, although the exact orientations of the primary gradients in the primary auditory cortex (A1) in HG are subject to debate (Saenz and Langers, 2014). We developed a population receptive field (pRF) model of the frequency-tuning response of individual voxels in the auditory cortex and compared center frequency (CF) and tuning sharpness (Q) in A1, rostral (R), and rostral-temporal (RT) areas among groups of AP musicians (AP group), matched non-AP musicians (MUS group), and controls without musical training (CON group). We hypothesized that the AP group would have sharper frequency tuning than the MUS and CON groups, which could explain the high accuracy of their pitch discrimination and identification (Bidelman et al., 2014).

## Materials and Methods

### Subjects

A total of 61 subjects were tested, including 20 AP [mean age ( $\pm$ SD)  $25.2 \pm 7.6$  years, 13 males], 20 MUS (mean age  $25.5 \pm 7.4$  years, 13 males), and 20 CON (mean age  $25.4 \pm 7.4$  years, 13 males) subjects. An additional musician subject scored high on the AP test but did not realize she had AP and was denoted as quasi-AP and omitted from the main groups. She reported using a tonal reference (middle C) and relative pitch comparisons on the AP test. AP and MUS subjects were recruited from notices advertised in university music departments and by word of mouth. Although AP is rare in the general population, AP subjects were readily identified; 18/20 of the AP subjects had music-related professions. A comprehensive auditory questionnaire was collected for each subject that pertained to musical background, education, primary instrument/voice, age of onset of musical training, and AP (Table 1). Subjects in each group were matched for age ( $F_{(2,59)} = 0.011, p = 0.99$ ), gender ( $F_{(2,59)} = 0, p = 1$ ), handedness ( $F_{(2,59)} = 0.38, p = 0.69$ ), and number of languages spoken ( $F_{(2,59)} = 0.66, p = 0.52$ ). Each group had three to four subjects who spoke a tonal language (e.g., Mandarin). AP and MUS subjects were matched on their primary instrument, onset age of musical training ( $F_{(1,39)} = 1.3, p = 0.27$ ), and the number of hours of musical training per week ( $F_{(1,39)} = 2.4, p = 0.44$ ). The ability to judge one note in relation to another given a reference tone is known as relative pitch (RP). RP is very common and all musicians (both AP and MUS) reported having RP. In the CON group, minimal musical training was defined as not having any current musical training in any instrument and having <3 years of any musical training and exposure overall. Of the 20 CON subjects, 11 had no musical training or exposure to any instrument, whereas 9 had minimal exposure, <3 years of musical training, and practiced <6 h/week during that time.

Before the collection of data, written informed consent was obtained from each subject after detailed explanation of the experimental procedure. All subjects were screened for normal hearing, had normal structural MRI scans, and did not report any hearing impairments or neurological disorders. The study was approved by the Human Participants Review Committee at York University.

### Behavioral tests

**AP test.** A standardized AP test (<http://www.musicianbrain.com/aptest/>), developed in the laboratory of Gottfried Schlaug, permitted objective classification of AP status. The AP test consisted of 24 sine wave tones drawn from the chromatic scale (C4–B4 repeated twice and randomized per trial). Data were collected on four trials (for a total of 96 tones presented). AP ability was confirmed if the accuracy within one semitone was 90% or above (Miyazaki, 1988; Zatorre and Beckett, 1989; Hamilton et al., 2004). Normal audiometric thresholds were confirmed for each subject using an audiometer.

**Table 1. Subject demographics and musical experience**

Group	AP	MUS	CON	Quasi-AP
No. of subjects	20	20	20	1
Sex (male/female)	13/7	13/7	13/7	Female
Age	22.2 (7.6)	25.5 (7.4)	25.4 (7.4)	22
Handedness				
Right-handed	17	15	16	
Left-handed	2	3	3	1
Ambidextrous	1	2	1	
Languages spoken				
Monolingual	11	10	9	
Bilingual	7	7	8	
Trilingual	2	3	3	
Tonal languages spoken	4	3	3	
Years of formal education	16.3 (2.1)	16.8 (2.8)	17.1 (2.2)	15
Age of musical training onset	5.7 (3.3)	6.8 (3.1)	–	6
Hours of practice per week	11.9 (7.6)	9.6 (5.1)	–	8
Primary instrument				
Piano	17	17	–	1
Guitar	2	2	–	
Trombone	1	1	–	

**Just noticeable difference (JND) test.** A JND frequency test was administered to determine the smallest detectable difference between two pitches. Two independent tests were run using different base frequencies: 1000 Hz (which does not correspond to a musical note) and 987.76 Hz (the equitempered tone of B5). The JND threshold for pitch depends on the frequency of the tone and sound level, as well as duration and the suddenness of the frequency change. We used a 250 ms pure tone with onsets and offsets ramped with a 10 ms cosine. The experiment was programmed in MATLAB (The MathWorks) using the Psychoacoustics Toolbox (Soranzo and Grassi, 2014). Two tones, the base frequency and a higher frequency, were presented in random order and the subjects pressed a key to indicate which pitch was higher. The higher tone was initially 100 Hz above the base frequency and changed on each trial based on the subjects' correct or incorrect responses according to a maximum likelihood procedure that efficiently estimated the difference threshold (Grassi and Soranzo, 2009). For each base frequency, 5 blocks were presented with 30 trials per block using a *p*-target (sweet point) = 80.9% and  $\beta = 0.5$  (slope of the logistic psychometric function).

**Melody mistuning (amusia) test.** To determine whether any of the controls had amusia, an additional behavioral test was administered pertaining to melody mistuning/tone deafness detection developed by Dr. Mandell (<http://jakemandell.com/tonedeaf/>). The test involved 36 short musical phrases that repeated twice. After hearing each repeated musical phrase, the participant would indicate if the melodies were both the same or different. The test was made purposefully difficult such that expert musicians on average scored 75% correct. Each melodic excerpt varied in musical timbre, duration, and tempo.

The behavioral tests were analyzed with SPSS (Mac version 23; IBM, RRID:SCR\_002865) and the  $\alpha$  value was Bonferroni adjusted for multiple comparisons for *post hoc* comparisons between groups.

### Imaging

All images were acquired using a 3 T Siemens Trio MRI scanner with a 32-channel head coil at York University. To reduce head motion, cushions were placed around the subjects' heads. A high-resolution T1-weighted 3D MPRAGE scan of the entire head was collected with the following parameters: TR = 1.9 s, TE = 2.52 ms, 1 mm thick slices,  $256 \times 256$  matrix (1 mm<sup>3</sup> isotropic voxel size). Ten functional runs with 160 time points each were collected in a single session with the following parameters: echo-planar, gradient echo sequence, 192 mm field of view, 128 matrix (1.5  $\times$  1.5  $\times$  2 mm voxel size), TR = 2 s, TE = 30 ms, 22.2 mm thick slices, flip angle = 90°, partial phase Fourier = 6/8, iPAT = GRAPPA, acceleration factor = 3. A whole head echo-planar image (WHEPI) was collected with the same parameters, except with 77 slices and TR = 7 s, to aid in registration.

### Stimuli

Auditory stimuli were presented on MRI-compatible on-ear piezo headphones (MR Confon). Subjects wore attenuating earplugs under the headphones and, to ensure a consistent balance between the background scanner noise and the stimulus, the volume settings of the computer and amplifier were fixed across subjects; all subjects reported clearly hearing the stimuli throughout the scanning procedures. Human hearing ranges from 20 Hz to 20 kHz, with the greatest sensitivity in the 200–2000 Hz range, which occupies up to two-thirds of the basilar membrane (Kollmeier et al., 2008). The stimuli consisted of pure tone logarithmic sweeps (chirps with unity amplitude and continuous phase) (De Martino et al., 2013) ranging from 20 Hz to 10 kHz, including 6 ascending and 6 descending 24 s sweeps in varying order interleaved with four 8 s blank periods. Subjects were instructed to fixate on a cross on the visual display and to remain still while listening to the auditory stimulus.

### Analysis

Data were preprocessed using MATLAB (RRID:SCR\_001622), FreeSurfer (RRID:SCR\_001847), and AFNI (RRID:SCR\_005927). Surface reconstructions were derived from using a single T1 collected at the beginning or end of each experimental scanning session. Cortical data were visualized in AFNI with the surface mapper SUMA. The anatomy was individually analyzed on surface inflated brains in the native space of each subject by a single rater blinded to group membership. HG was readily apparent in this space and parcellations of A1, R, and RT regions of interest (ROIs) were based on the anatomy and directions of tonotopic gradients. For subjects with a bifurcated and trifurcated HG, the most anterior portion was chosen using low-to-high gradients as landmarks, consistent with previous studies (Kaas and Hackett, 2000; Humphries et al., 2010; Moerel et al., 2014). Anatomical statistics were calculated using SPSS (Mac version 25; IBM, RRID:SCR\_002865).

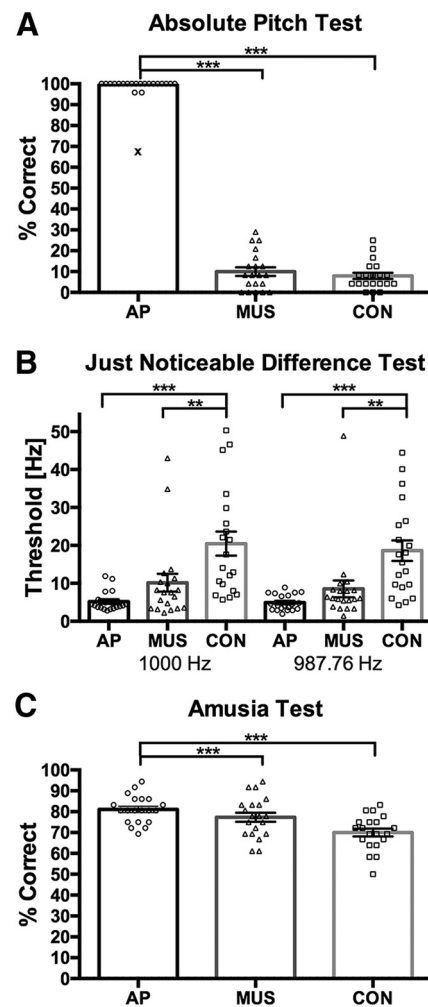
Each functional run was aligned by composing a two-step affine registration of the EPI slab registered to the WHEPI and the WHEPI registered to the surface anatomy. The first four volumes of each functional scanning run were discarded. To compensate for subject head movement, the remaining volumes were registered to a single volume obtained during the same scanning session. In addition to motion correcting the functional imaging data, we also used a volume-censoring procedure (Power et al., 2012). The framewise displacement is an aggregate measure of the translational and rotational head movement gleaned from the motion correction transformation for each volume in the functional series. Volumes with a framewise displacement  $>0.4$  mm were flagged for censoring and were not included in the mean functional series. The functional data were time shifted and de-obliqued and the linear trend was removed from the time series of each voxel. The functional data were resampled to  $0.75 \text{ mm}^3$  and smoothed in surface space with a 3 mm full width at half maximum (FWHM) kernel and combined into a single mean run.

Data were analyzed using an open source (DeSimone et al., 2016) adaptation of the pRF model (Dumoulin and Wandell, 2008). pRF maps were thresholded at  $r^2 = 0.0625$  ( $r = 0.25$ ),  $p = 0.01$ . We used a 1D Gaussian pRF model to describe the BOLD response of each voxel in terms of stimulus-referred CF and bandwidth (BW), which was the SD of the Gaussian. The model fitting consisted of two phases: a brute force grid search of a sparsely sampled stimulus space to find the initial guess and a subsequent gradient descent search in a more finely sampled space. In addition to the two parameters of the Gaussian frequency response, we also modeled the baseline and amplitude of the voxelwise BOLD signal to account for its arbitrary units. We allowed the model to explore the frequency space 2 Hz–10 kHz. Q was computed as the ratio between CF and the FWHM of the Gaussian pRF response model, where  $\text{FWHM} = 2.355 * \text{SD}$ . Therefore, sharper tuning relates to a higher Q value, whereas broader tuning relates to a lower Q value.

## Results

### Behavioral measures

A one-way ANOVA revealed a significant main effect of group ( $F_{(2,59)} = 1248.6$ ,  $p < 0.001$ ) among the AP, MUS, and CON groups for the AP test. *Post hoc* tests between groups revealed



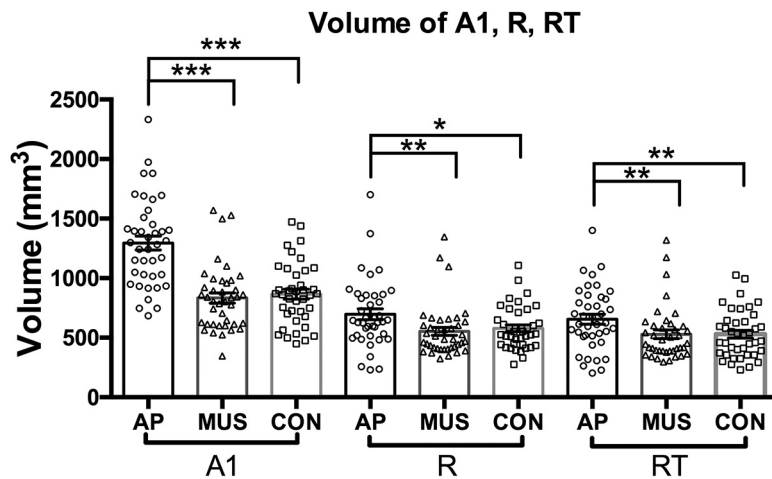
**Figure 1.** Behavioral test scores. AP: open black circles, MUS: open triangles, CON: open squares.  $n = 20$  per group. **A**, Absolute pitch test scores. “X” represents the quasi-AP subject who scored 67% correct on the AP test. **B**, JND thresholds. **C**, Melody mistuning detection (amusia) test results. Error bars indicate SEM. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

significantly increased scores for AP ( $99.58 \pm 0.29\%$ , mean  $\pm$  SEM) compared with both MUS ( $10.0 \pm 9.1\%$ ,  $p < 0.001$ ) and CON ( $7.9 \pm 6.7\%$ ,  $p < 0.001$ ), but no significant difference between MUS and CON ( $p = 0.33$ ) (Fig. 1).

JNDs were submitted to a one-way repeated-measures ANOVA with base frequency (1000 or 987.76 Hz) as the repeated measure. A significant main effect of group ( $F_{(2,57)} = 13.0$ ,  $p < 0.001$ ) was observed but no main effect of frequency ( $F_{(1,57)} = 2.6$ ,  $p = 0.11$ ) or interaction between group and frequency ( $F_{(2,57)} = 0.4$ ,  $p = 0.67$ ). Comparisons between groups showed that both AP ( $5.1 \pm 2.1$  Hz) and MUS ( $9.3 \pm 2.1$  Hz) showed smaller JNDs than CON ( $19.6 \pm 2.1$  Hz,  $p < 0.001$  and  $p = 0.001$ , respectively), but did not significantly differ from each other ( $p = 0.15$ ). Of particular interest was whether the AP subjects would differ on the JND test between frequencies because only one of the frequencies corresponded to a named musical note, but a paired-samples  $t$  test revealed no significant effect of frequency on JND in this group ( $t_{(19)} = 0.53$ ,  $p = 0.61$ ) (Fig. 1B).

A one-way ANOVA revealed a significant main effect of group ( $F_{(1,59)} = 9.27$ ,  $p < 0.001$ ) for the melody mistuning test. *Post hoc* tests revealed significantly increased scores for AP compared with CON ( $p < 0.001$ ) and MUS compared with CON ( $p = 0.021$ ). There were no significant differences between AP and MUS ( $p =$





**Figure 2.** Volumes of A1, R, and RT, collapsed across hemispheres. AP: open black circles, MUS: open black triangles, CON: open black squares.  $n = 20$  per group. Error bars indicate SEM. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

0.47). The AP test average score in AP participants was  $81.1 \pm 1.4\%$ ; in MUS, it was  $77.4 \pm 2.1\%$ ; and in CON, it was  $70.0 \pm 1.8\%$  (Fig. 1C). CON subjects ranged in test scores from 50% to 83% correct and were within the distribution of  $>61,000$  participants who took the test online.

## Imaging

### Anatomy

A total of 58.3% of all subjects collapsed across hemispheres and groups had an intact HG, whereas 31.7% had their HG partially bifurcate and 10% had their HG partially trifurcate. Interestingly, the left hemisphere had fewer bifurcations or trifurcations (30%) than the right hemisphere (53.3%). We subjected the total volume of HG (A1 + R + RT) to a one-way repeated-measures ANOVA with hemisphere as the within-subjects factor. We found a significant main effect of group ( $F_{(2,57)} = 10.7$ ,  $p < 0.001$ ). Planned comparisons revealed that the HG volume in the AP ( $2661 \pm 127 \text{ mm}^3$ ) was significantly larger than in either MUS ( $1915 \pm 127 \text{ mm}^3$ ,  $p < 0.001$ ) or CON ( $1973 \pm 127 \text{ mm}^3$ ,  $p < 0.001$ ), which did not differ significantly from each other ( $p = 0.75$ ). In addition, there was a significant main effect of hemisphere ( $F_{(1,57)} = 9.6$ ,  $p = 0.003$ ) with the left HG ( $2326 \pm 104 \text{ mm}^3$ ) significantly larger than the right ( $2041 \pm 64 \text{ mm}^3$ ), but there was no significant interaction between group and hemisphere ( $F_{(2,57)} = 0.38$ ,  $p = 0.69$ ). A Pearson product-moment correlation coefficient (two-tailed) was computed to assess the relationship between the volume of HG in the AP and MUS groups and the mean number of practice hours per week and the onset of age of musical training, but no significant correlations were found.

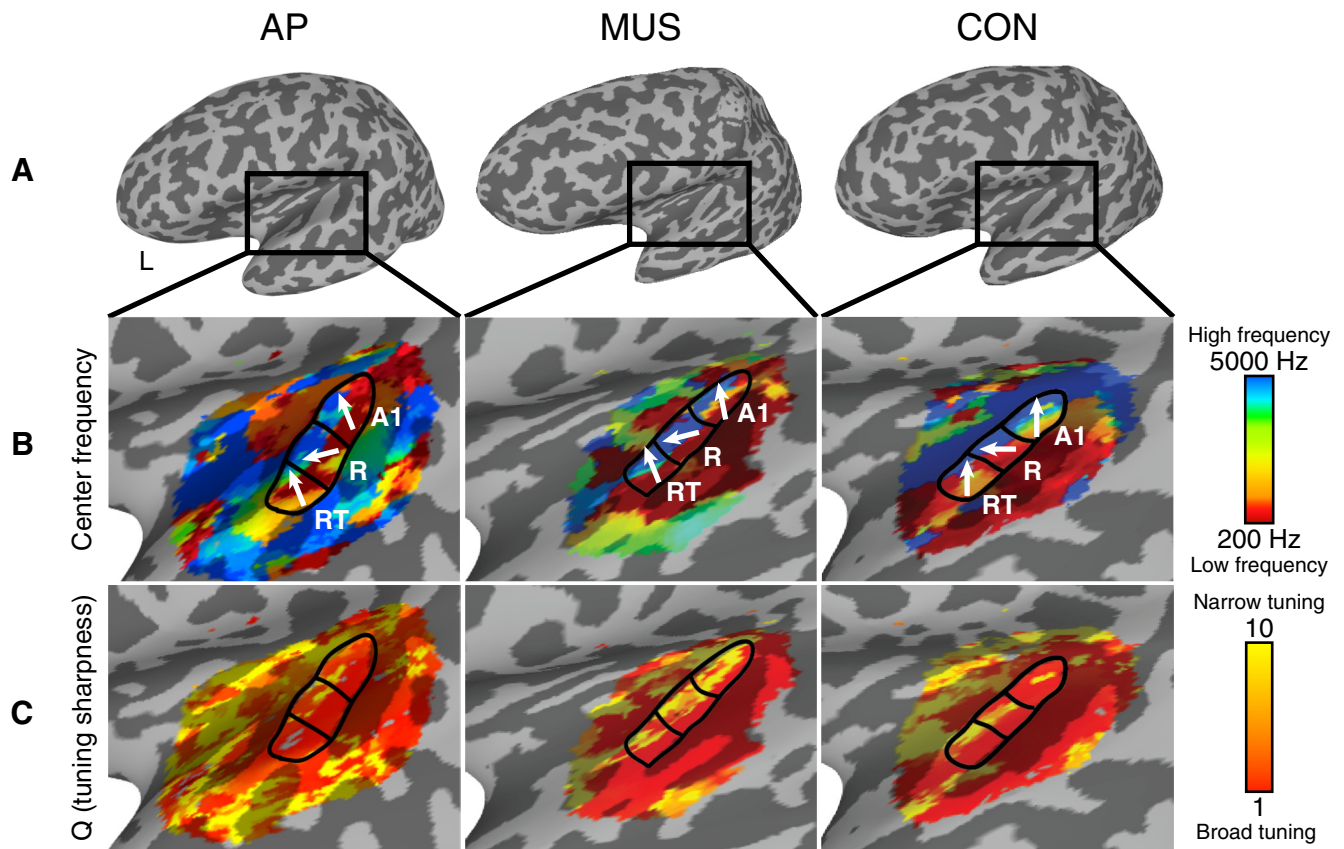
The volumes of the individual cortical areas A1, R, and RT were each subjected to a one-way repeated-measures ANOVA, with hemisphere as the repeated measure (Fig. 2). For A1, there was a significant main effect of group ( $F_{(2,57)} = 21.8$ ,  $p < 0.001$ ). Planned comparisons revealed that A1 volume was greater ( $p < 0.001$ ) in AP ( $1284 \pm 53 \text{ mm}^3$ ) than in MUS ( $832 \pm 53 \text{ mm}^3$ ) or CON ( $870 \pm 53 \text{ mm}^3$ ), but did not significantly differ between MUS and CON ( $p = 0.63$ ). The main effect of hemisphere was also significant ( $F_{(1,57)} = 16.6$ ,  $p < 0.001$ ), with the left hemisphere larger than the right,  $1068 \pm 41$  vs  $922 \pm 29 \text{ mm}^3$ , but no significant interaction between group and hemisphere ( $F_{(2,57)} = 0.49$ ,  $p = 0.61$ ). For R, there was a significant main effect of group ( $F_{(2,57)} = 5.12$ ,  $p = 0.009$ ). Planned comparisons revealed that R

volume was greater in AP ( $719 \pm 40 \text{ mm}^3$ ) than in MUS ( $549 \pm 40 \text{ mm}^3$ ,  $p = 0.004$ ) or CON ( $579 \pm 40 \text{ mm}^3$ ,  $p = 0.016$ ), but did not significantly differ between MUS and CON ( $p = 0.61$ ). The main effect of hemisphere was also significant ( $F_{(1,57)} = 4.57$ ,  $p = 0.037$ ), with the left hemisphere larger than the right,  $655 \pm 34$  vs  $576 \pm 24 \text{ mm}^3$ , but no significant interaction between group and hemisphere ( $F_{(2,57)} = 0.19$ ,  $p = 0.83$ ). For RT, there was a significant main effect of group ( $F_{(2,57)} = 6.08$ ,  $p = 0.004$ ). Planned comparisons revealed that RT volume was greater in AP ( $690 \pm 40 \text{ mm}^3$ ) than in MUS ( $530 \pm 40 \text{ mm}^3$ ,  $p = 0.006$ ) or CON ( $515 \pm 40 \text{ mm}^3$ ,  $p = 0.003$ ), but did not significantly differ between MUS and CON ( $p = 0.78$ ). The main effect of hemisphere was only marginally significant ( $F_{(1,57)} = 2.91$ ,  $p = 0.093$ ), with the left hemisphere larger than the right,  $607 \pm 32$  vs  $549 \pm 25 \text{ mm}^3$ , and there was no significant interaction between group and hemisphere ( $F_{(2,57)} = 0.35$ ,  $p = 0.70$ ).

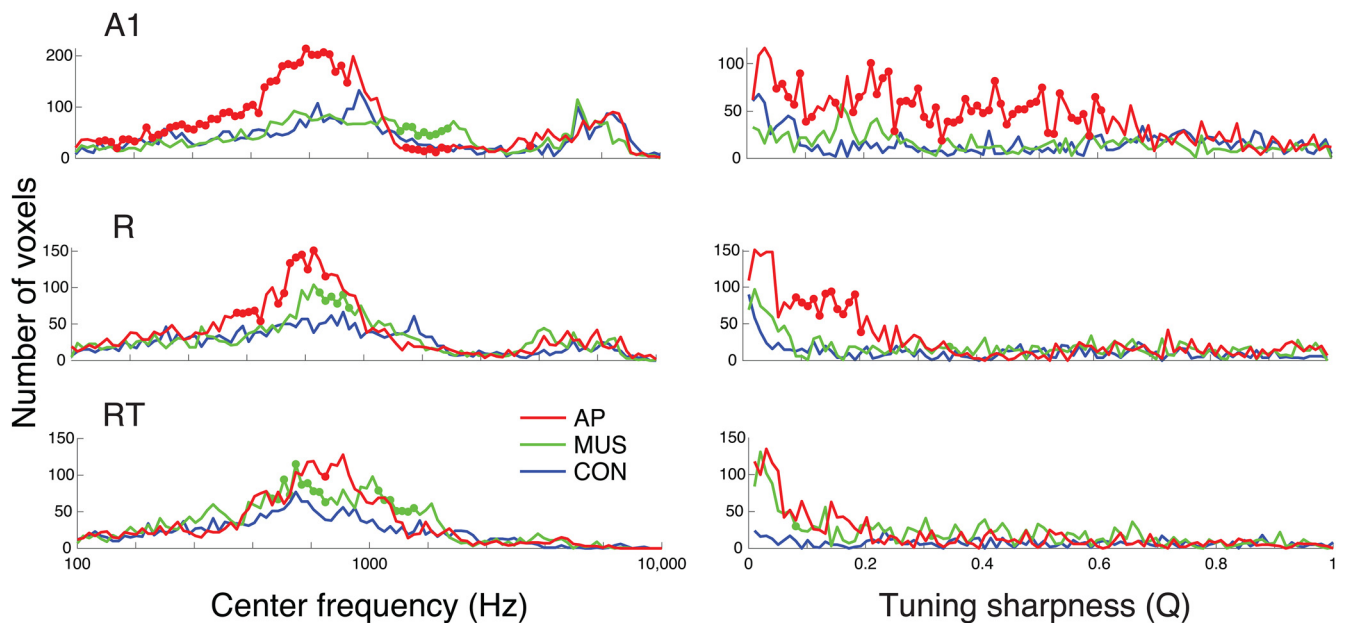
### Function

We conducted a one-way ANOVA comparing the percentage of time points censored for excessive motion. Across groups,  $4.32 \pm 0.70\%$  of time points were removed due to a framewise displacement of  $>0.4 \text{ mm}$ , but this percentage did not significantly vary among the three groups ( $F_{(2,59)} = 0.074$ ,  $p = 0.93$ ). The frequency tuning of voxels in the three cortical areas were generally well described by the Gaussian pRF model. Across all subjects and areas, 33.2% of voxels exhibited below threshold responses to the auditory stimuli. In 34.0% of the remaining activated voxels, the fitted CF and/or BW parameters fell outside 100 Hz–10 kHz. Therefore, of the voxels located within the anatomically defined auditory areas, 44.1% were well activated and well characterized by the Gaussian pRF model and their parameters were subjected to further analysis. To confirm that there were no biases in activation across groups, the percentage of activated voxels in each area were subjected to a repeated-measures ANOVA, with hemisphere and area as within-subjects factors. The main effect of group was not significant ( $F_{(1,57)} = 2.06$ ,  $p = 0.14$ ), but there was a significant effect of hemisphere ( $F_{(1,57)} = 4.40$ ,  $p = 0.040$ ), with the right hemisphere better activated than the left ( $46.2 \pm 2.6$  vs  $41.8 \pm 2.6\%$ ). There was also a significant interaction between hemisphere and area ( $F_{(1,57)} = 6.91$ ,  $p = 0.011$ ), with the asymmetry most pronounced in area RT, with  $49.0 \pm 3.6\%$  of voxels activated in the right hemisphere compared with  $39.4 \pm 3.5\%$  in the left.

Examining the topography of CF representation, we found consistent tonotopic maps with a high-to-low frequency gradient in A1, reversing to low-to-high frequency in R and reversing again to high-to-low in RT; the gradients extended into surrounding areas (Moerel et al., 2014). However, our results may also be interpreted as parallel high-to-low gradients within each auditory region (Saenz and Langers, 2014). The topography of Q was less well defined but approximately followed the gradient of CF, with narrow-to-broad tuning oriented in the same direction as high-to-low frequency. Typical examples for each group of the distributions over the early auditory cortex of CF and Q are shown in Figure 3.



**Figure 3.** pRF maps of tonotopy (center frequency) and Q in auditory cortex in representative subjects from each group. **A**, Left surface inflated hemispheres. **B**, Left zoomed in view of the pRF map for A1, R, RT and belt regions. Solid black lines indicate boundaries between tonotopic maps and black arrows indicate direction of tonotopic gradient (low-high) consistent with previous interpretations (Kaas and Hackett, 2000; Moerel et al., 2014). **C**, Left magnified view of the Q maps.



**Figure 4.** Center frequency and Q distributions for A1, R, and RT in the AP group (red), the MUS group (green), and the CON group (blue). Each distribution is across all subjects in that group. The red dots indicate points at which the AP distribution differs significantly from the MUS distribution and the green dots indicate points at which the MUS distribution differs significantly from CON.

To qualitatively compare the auditory representations among the three groups, we examined the distribution of CF and Q (Fig. 4). Voxels were pooled from all subjects and hemispheres in each group and CF and Q were sorted into 100 bins, ranging from 100

to 10,000 Hz and  $Q = 0-1$ , respectively. We measured the statistical significance of the difference at each bin in the distributions by calculating a paired *t* test that included 5 bins on either side, adjusted for multiple comparisons (100 bins) using a reduced

$\alpha = 0.0005$ . Planned comparisons were made between the AP and MUS groups and between the MUS and CON groups. In Figure 4, the bins in the distribution that were significantly different from each other are indicated by circular markers (red for AP vs MUS and green for MUS vs CON). As can be seen, the increase in cortical area in A1 for the AP group compared with MUS was mostly due to a larger number of voxels with CF < 1000 Hz and those with broad tuning,  $Q < 0.6$ . Interestingly, the AP group had significantly fewer voxels with CF in the 1250–2000 Hz range compared with the MUS group, which in turn had significantly more voxels in this range than the CON group. Similar, but smaller differences between AP and MUS were observed in area R for CF in the range of 400–800 Hz and for  $Q = 0.1–0.2$ . There were few differences between AP and MUS in area RT for either CF or  $Q$ . There were few differences at all in  $Q$  between the MUS and CON groups in any area, but the MUS group had a significantly larger number of voxels than the CON group with CF ~500–1000 Hz in areas R and RT.

## Discussion

This has been the first study to measure the tonotopic and tuning sharpness organization in a large sample size of the AP, MUS, and CON groups within the A1, R, and RT subdivisions of HG. We observed that A1, R, and RT were significantly larger in AP than MUS or CON subjects, which did not differ significantly from each other. The increased A1 and R area in AP was primarily due to increases in the volume devoted to low-frequency and broadly tuned responses, whereas the distribution of tuning responses in area RT did not differ significantly.

Previous studies in nonhuman primates found sharper (narrower) neural tuning within A1 core regions and broader tuning in belt regions (Rauschecker et al., 1995; Hackett et al., 1998; Rauschecker and Tian, 2004; Kajikawa et al., 2005; Kuśmierk and Rauschecker, 2009). Tuning widths were also reported narrower in the human auditory core regions compared with nonhuman primates based on electrophysiological recordings (Bitterman et al., 2008) and a filter model of cochlear responses suggested sharper tuning in musicians (Bidelman et al., 2014). Whereas broader frequency tuning might be counterintuitive, it is suggestive of a greater utilization of ensemble encoding of frequency, with more cortical machinery involved in the encoding of frequencies in AP subjects. Ensemble encoding can allow more efficient and thus higher bandwidth representations (Cohen et al., 2016). Additionally, frequency tuning within a voxel in AP subjects may be broader either due to the underlying neurons themselves having broader tuning or due to more scatter among the neurons within the voxel. Given that there are smooth tonotopic gradients in auditory cortex, the former seems more likely.

We found consistent tonotopic map interpretations of the cortex that matched previous studies: a high-to-low progression in A1, followed by a reversal gradient of low-to-high in R, followed by a gradient of high-to-low in RT, with extended gradients into neighboring belt regions as found in neuroimaging studies in humans (Da Costa et al., 2011; Moerel et al., 2012, 2014; De Martino et al., 2013) and in microelectrode studies nonhuman primates (Morel et al., 1993; Kaas and Hackett, 2000). The auditory cortical anatomical model in the monkey has been well defined predominantly based on neuro-electrical recordings (Kaas, 2011). However, there exist many differences in the human compared with monkey auditory cortex, including larger cortical surface areas, additional gyri, more interindividual variability, and sharper frequency tuning in humans (Galaburda et al., 1978; Bitterman et al., 2008). Therefore, it may not be

straightforward to apply the monkey model to the human brain for direct comparison. A variety of parcellation schemes have been proposed for the human auditory cortex. Barton et al. (2012) suggest a cloverleaf parcellation of auditory cortex based on their periodotopy measurements. However, Herdener et al. (2013) suggested a different parcellation using periodotopy, and Leaver and Rauschecker (2016) found no evidence to support large-scale organized periodotopy at all. For further review of the inconsistencies in parcellation schemes, see Moerel et al. (2014). Our results are not strongly dependent on the precise parcellation scheme because we observed a broad and gradually decreasing trend from A1 to R to RT.

Based on our behavioral tests in this study, we report that AP and MUS had significantly smaller JND thresholds than their CON counterparts. Our findings are consistent with a previous study finding no differences in JND thresholds between AP and MUS groups (Fujisaki and Kashino, 2002), although musicians overall have done significantly better than non-musician controls (Micheyl et al., 2006). In addition, a previous study found sharper cochlear tuning in a high 4 kHz frequency region in musicians compared with non-musicians using both forward- and simultaneous-masked psychophysical tuning curves and a relationship between years of musical training and sharper tuning (Bidelman et al., 2014).

An ancillary finding to this study was that 20% of the AP subjects who were recruited did not have any musical training before the age of 7 years (the critical period window) and only started any formal musical training and note association labeling in their early to late teens. This finding does not support the critical period hypothesis suggesting that a child must be exposed to musical training in note labeling before the age of seven for AP ability to emerge (Levitin and Zatorre, 2003; Russo et al., 2003; Miyazaki and Ogawa, 2006). Although debated, further claims that AP is linked to a critical period suggest that musical training after the age of 9 years very seldomly leads to AP emergence, which is additionally supported by no reported cases in adults successfully developing it (Brady, 1970; Cohen and Baird, 1990; Ward, 1999; Levitin and Rogers, 2005). Our findings suggest that genetics may play a more salient role for AP ability to emerge in neurodevelopment as opposed to a critical period alone.

To our knowledge, this is the first study to separately extract the volume of auditory ROIs comprising HG (A1, R, and RT) in humans categorized by pitch perception attributes. Of most significance, A1 volume was significantly larger in both hemispheres in AP compared with MUS and CON subjects. Our results are consistent with reported findings that A1 occupies approximately half of the HG volume (Rademacher et al., 2001). Within A1, neurons exhibit characteristic responses to harmonic spectral stimuli and periodic temporal modulations (Wang, 2013). Rats with bilateral A1 inactivation showed impairments in their response to pure tone frequency changes (Talwar et al., 2001). It seems that A1 does have some related function with auditory pitch discrimination. However, it is not clear whether only a subset of A1 neurons participates in pitch encoding, with other neurons analyzing temporal or spectral components of sound or if pitch is more preferentially encoded in secondary cortical fields beyond A1. Nonetheless, A1 is implicated in AP, suggesting that its enhanced volume may be related to AP ability in pitch categorization and perception.

Areas outside of the auditory core, including the belt, parabelt, and regions beyond, may play a more relevant role in pitch perception. For example, a previous fMRI study reported cortical activation in response to pitch height that extended beyond au-



ditary core regions into the posterior planum temporale, whereas cortical activation in response to pitch chroma (i.e., pitch class, where a set of pitches are related to each other by octave) extended to the planum polare, a region just anterior to A1; a hierarchical stream of pitch processing was proposed to account for these findings, with areas beyond the primary auditory cortex having specialized perceptual roles (Warren et al., 2003). More anterior regions to the core were responsive to object-independent auditory stimuli, whereas more posterior regions to the core including the planum temporale were more responsive to object identification. Future studies need to account for these extended regions involved in pitch processing, including the read-out of lower-level representations, which we found to be markedly different in AP compared with MUS and CON subjects.

We report all auditory ROIs in left HG trended larger than those in the right HG, with A1 and R being significantly larger. Several studies have reported that left HG was ~10–30% larger than the right using MRI (Penhune et al., 1996; McCarley et al., 2002; Sumich et al., 2002; Dorsaint-Pierre et al., 2006; Takahashi et al., 2006; Golestani et al., 2007; Salisbury et al., 2007) and postmortem measurements (Chance et al., 2008; Smiley et al., 2013), whereas other neuroimaging studies, including those with large sample sizes, did not report hemispheric asymmetry in HG (Kulynych et al., 1995; Frangou et al., 1997; Schneider et al., 2002; Knaus et al., 2006). We used the stringent boundary delineations of the recent working model of the human auditory cortex that only includes A1, R, and RT in HG (Moerel et al., 2014). The reported discrepancies on asymmetry in HG are likely due to various interpretations and methods of defining the borders of HG. A few studies included surrounding areas of the planum temporale and planum polare within HG, which lead to considerable overestimation of the auditory core size and potentially biased the asymmetry measurements (Da Costa et al., 2011; Herdener et al., 2013; Langers, 2014).

In conclusion, we found that each of the auditory areas in HG was significantly larger in subjects with AP compared to MUS and CON groups, and that this increased cortical area was primarily broadly tuned to frequencies below 1000 Hz.

## References

- Bachem A (1955) Absolute pitch. *J Acoust Soc Am* 27:1180–1185.
- Barton B, Venezia JH, Saberi K, Hickok G, Brewer AA (2012) Orthogonal acoustic dimensions define auditory field maps in human cortex. *Proc Natl Acad U S A* 109:20738–20743.
- Bidelman GM, Schug JM, Jennings SG, Bhagat SP (2014) Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians. *J Acoust Soc Am* 136:EL33–EL39.
- Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature* 451:197–201.
- Brady PT (1970) Fixed-scale mechanism of absolute pitch. *J Acoust Soc Am* 48:883–887.
- Chance SA, Casanova MF, Switala AE, Crow TJ (2008) Auditory cortex asymmetry, altered minicolumn spacing and absence of ageing effects in schizophrenia. *Brain* 131:3178–3192.
- Cohen AJ, Baird K (1990) Acquisition of absolute pitch: the question of critical period. *Psychomusicology* 9:31–37.
- Cohen MA, Dennett DC, Kanwisher N (2016) What is the bandwidth of perceptual experience? *Trends Cogn Sci* 20:324–335.
- Da Costa S, van der Zwaag W, Marques JP, Frackowiak RS, Clarke S, Saenz M (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci* 31:14067–14075.
- De Martino F, Moerel M, van de Moortele PF, Ugurbil K, Goebel R, Yacoub E, Formisano E (2013) Spatial organization of frequency preference and selectivity in the human inferior colliculus. *Nat Commun* 4:1386.
- DeSimone K, Rokem A, Schneider K (2016) popeye: a population receptive field estimation tool. *J Open Source Softw* 1:103.
- Deutsch D (2013) Absolute pitch. In: *The psychology of music*, Ed 3 (Deutsch D, ed), pp 141–182. San Diego: Elsevier, Inc.
- Deutsch D, Henthorn T, Marvin E, Xu H (2006) Absolute pitch among american and chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *J Acoust Soc Am* 119:719–722.
- Deutsch D, Dooley K, Henthorn T, Head B (2009) Absolute pitch among students in an american music conservatory: association with tone language fluency. *J Acoust Soc Am* 125:2398–2403.
- Dohn A, Garza-Villarreal EA, Chakravarty MM, Hansen M, Lerch JP, Vuust P (2015) Gray- and white-matter anatomy of absolute pitch possessors. *Cereb Cortex* 25:1379–1388.
- Dorsaint-Pierre R, Penhune VB, Watkins KE, Neelin P, Lerch JP, Bouffard M, Zatorre RJ (2006) Asymmetries of the planum temporale and Heschl's gyrus: relationship to language lateralization. *Brain* 129:1164–1176.
- Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. *Neuroimage* 39:647–660.
- Frangou S, Sharma T, Sigmundsson T, Barta P, Pearson G, Murray RM (1997) The maudslay family study. 4. normal planum temporale asymmetry in familial schizophrenia: a volumetric MRI study. *Br J Psychiatry* 170:328–333.
- Fujisaki W, Kashino M (2002) The basic hearing abilities of absolute pitch possessors. *Acoust Sci Technol* 23:77–83.
- Galaburda AM, LeMay M, Kemper TL, Geschwind N (1978) Right-left asymmetries in the brain. *Science* 199:852–856.
- Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C (2007) Brain structure predicts the learning of foreign speech sounds. *Cereb Cortex* 17:575–582.
- Grassi M, Soranzo A (2009) MLP: a MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behav Res Methods* 41:20–28.
- Hackett TA, Stepniewska I, Kaas JH (1998) Thalamocortical connections of the parabelt auditory cortex in macaque monkeys. *J Comp Neurol* 400:271–286.
- Hamilton RH, Pascual-Leone A, Schlaug G (2004) Absolute pitch in blind musicians. *Neuroreport* 15:803–806.
- Herdener M, Esposito F, Scheffler K, Schneider P, Logothetis NK, Uludag K, Kayser C (2013) Spatial representations of temporal and spectral sound cues in human auditory cortex. *Cortex* 49:2822–2833.
- Humphries C, Liebenthal E, Binder JR (2010) Tonotopic organization of human auditory cortex. *Neuroimage* 50:1202–1211.
- Kaas JH (2011) The evolution of auditory cortex: The core areas. In: *The auditory cortex* (Winer, Jeffery A., Schreiner, Christoph E., Eds.), pp 407–427. New York: Springer.
- Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A* 97:11793–11799.
- Kajikawa Y, de La Mothe L, Blumell S, Hackett TA (2005) A comparison of neuron response properties in areas A1 and CM of the marmoset monkey auditory cortex: tones and broadband noise. *J Neurophysiol* 93:22–34.
- Knaus TA, Bollich AM, Corey DM, Lemen LC, Foundas AL (2006) Variability in perisylvian brain anatomy in healthy adults. *Brain Lang* 97:219–232.
- Kollmeier B, Brand T, Meyer B (2008) Perception of Speech and Sound. In: *Springer handbook of speech processing* (Benesty J, Sondhi MM, Huang Y, eds), pp 61–82. Berlin: Springer.
- Kulynych JJ, Vldar K, Fantie BD, Jones DW, Weinberger DR (1995) Normal asymmetry of the planum temporale in patients with schizophrenia: three-dimensional cortical morphometry with MRI. *Br J Psychiatry* 166:742–749.
- Kuśmerek P, Rauschecker JP (2009) Functional specialization of medial auditory belt cortex in the alert rhesus monkey. *J Neurophysiol* 102:1606–1622.
- Langers DR (2014) Assessment of tonotopically organised subdivisions in human auditory cortex using volumetric and surface-based cortical alignments. *Hum Brain Mapp* 35:1544–1561.
- Leaver AM, Rauschecker JP (2016) Functional topography of human auditory cortex. *J Neurosci* 36:1416–1428.
- Levitin DJ, Rogers SE (2005) Absolute pitch: perception, coding, and controversies. *Trends Cogn Sci* 9:26–33.
- Levitin DJ, Zatorre RJ (2003) On the nature of early music training and absolute pitch: a reply to brown, Sachs, Cammuso, and Folstein. *Music Percept* 21:105–110.
- Loui P, Li HC, Hohmann A, Schlaug G (2011) Enhanced cortical connec-

- tivity in absolute pitch musicians: a model for local hyperconnectivity. *J Cogn Neurosci* 23:1015–1026.
- Loui P, Zamm A, Schlaug G (2012) Enhanced functional networks in absolute pitch. *Neuroimage* 63:632–640.
- McCarley RW, Salisbury DF, Hirayasu Y, Yurgelun-Todd DA, Tohen M, Zarate C, Kikinis R, Jolesz FA, Shenton ME (2002) Association between smaller left posterior superior temporal gyrus volume on magnetic resonance imaging and smaller left temporal P300 amplitude in first-episode schizophrenia. *Arch Gen Psychiatry* 59:321–331.
- Michéyl C, Delhommeau K, Perrot X, Oxenham AJ (2006) Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res* 219:36–47.
- Miyazaki K (1988) Musical pitch identification by absolute pitch possessors. *Percept Psychophys* 44:501–512.
- Miyazaki K, Ogawa Y (2006) Learning absolute pitch by children. *Music Percept* 24:63–78.
- Moerel M, De Martino F, Formisano E (2012) Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. *J Neurosci* 32:14205–14216.
- Moerel M, De Martino F, Formisano E (2014) An anatomical and functional topography of human auditory cortical areas. *Front Neurosci* 8:225.
- Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol* 335:437–459.
- Penhune VB, Zatorre RJ, MacDonald JD, Evans AC (1996) Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb Cortex* 6:661–672.
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59:2142–2154.
- Profita J, Bidder TG (1988) Perfect pitch. *Am J Med Genet* 29:763–771.
- Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage* 13:669–683.
- Rauschecker JP, Tian B (2004) Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 91:2578–2589.
- Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.
- Russo FA, Windell DL, Cuddy LL (2003) Learning the “special note”: evidence for a critical period for absolute pitch acquisition. *Music Percept* 21:119–127.
- Sacks OW (2007) *Musicophilia: Tales of music and the brain*. New York: Random House.
- Saenz M, Langers DR (2014) Tonotopic mapping of human auditory cortex. *Hear Res* 307:42–52.
- Salisbury DF, Kuroki N, Kasai K, Shenton ME, McCarley RW (2007) Progressive and interrelated functional and structural evidence of post-onset brain reduction in schizophrenia. *Arch Gen Psychiatry* 64:521–529.
- Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A (2002) Morphology of Heschl’s gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci* 5:688–694.
- Schulze K, Gaab N, Schlaug G (2009) Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neurosci* 10:106.
- Smiley JF, Hackett TA, Preuss TM, Bleiwas C, Figarsky K, Mann JJ, Rosoklija G, Javitt DC, Dwork AJ (2013) Hemispheric asymmetry of primary auditory cortex and Heschl’s gyrus in schizophrenia and nonpsychiatric brains. *Psychiatry Res* 214:435–443.
- Soranzo A, Grassi M (2014) PSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing. *Front Psychol* 5:712.
- Sumich A, Chitnis XA, Fannon DG, O’Ceallaigh S, Doku VC, Falrowicz A, Marshall N, Matthew VM, Potter M, Sharma T (2002) Temporal lobe abnormalities in first-episode psychosis. *Am J Psychiatry* 159:1232–1235.
- Takahashi T, Suzuki M, Zhou SY, Tanino R, Hagino H, Kawasaki Y, Matsui M, Seto H, Kurachi M (2006) Morphologic alterations of the parcellated superior temporal gyrus in schizophrenia spectrum. *Schizophr Res* 83:131–143.
- Talwar SK, Musial PG, Gerstein GL (2001) Role of mammalian auditory cortex in the perception of elementary sound properties. *J Neurophysiol* 85:2350–2358.
- Wang X (2013) The harmonic organization of auditory cortex. *Front Syst Neurosci* 7:114.
- Ward WD (1999) Absolute pitch. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed.) (pp. 265–298) New York: Academic Press.
- Warren JD, Uppenkamp S, Patterson RD, Griffiths TD (2003) Analyzing pitch chroma and pitch height in the human brain. *Ann N Y Acad Sci* 999:212–214.
- Zatorre RJ, Beckett C (1989) Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. *Mem Cognit* 17:582–589.