

Pitch-Responsive Cortical Regions in Congenital Amusia

Sam V. Norman-Haignere,¹ Philippe Albouy,^{3,4} Anne Caclin,³ Josh H. McDermott,¹ Nancy G. Kanwisher,^{1,2} and Barbara Tillmann³

¹Department of Brain and Cognitive Sciences and ²McGovern Institute for Brain Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, ³Lyon Neuroscience Research Centre, Centre National de la Recherche Scientifique UMR5292, INSERM U1028, Lyon 1 University, F-69000 Lyon, France, and ⁴Montreal Neurological Institute, McGill University, Quebec H3A 0G4, Canada

Congenital amusia is a lifelong deficit in music perception thought to reflect an underlying impairment in the perception and memory of pitch. The neural basis of amusic impairments is actively debated. Some prior studies have suggested that amusia stems from impaired connectivity between auditory and frontal cortex. However, it remains possible that impairments in pitch coding within auditory cortex also contribute to the disorder, in part because prior studies have not measured responses from the cortical regions most implicated in pitch perception in normal individuals. We addressed this question by measuring fMRI responses in 11 subjects with amusia and 11 age- and education-matched controls to a stimulus contrast that reliably identifies pitch-responsive regions in normal individuals: harmonic tones versus frequency-matched noise. Our findings demonstrate that amusic individuals with a substantial pitch perception deficit exhibit clusters of pitch-responsive voxels that are comparable in extent, selectivity, and anatomical location to those of control participants. We discuss possible explanations for why amusics might be impaired at perceiving pitch relations despite exhibiting normal fMRI responses to pitch in their auditory cortex: (1) individual neurons within the pitch-responsive region might exhibit abnormal tuning or temporal coding not detectable with fMRI, (2) anatomical tracts that link pitch-responsive regions to other brain areas (e.g., frontal cortex) might be altered, and (3) cortical regions outside of pitch-responsive cortex might be abnormal. The ability to identify pitch-responsive regions in individual amusic subjects will make it possible to ask more precise questions about their role in amusia in future work.

Key words: amusia; auditory cortex; fMRI; music; pitch

Significance Statement

The neural causes of congenital amusia, a lifelong deficit in pitch and music perception, are not fully understood. We tested the hypothesis that amusia is due to abnormalities in brain regions that respond selectively to sounds with a pitch in normal listeners. Surprisingly, amusic individuals exhibited pitch-responsive regions that were similar to normal-hearing controls in extent, selectivity, and anatomical location. We discuss how our results inform current debates on the neural basis of amusia and how the ability to identify pitch-responsive regions in amusic subjects will make it possible to ask more precise questions about their role in amusic deficits.

Introduction

Individuals with congenital amusia experience difficulty performing simple musical tasks, such as recognizing a melody or detecting an out-of-key note (Peretz et al., 2002), despite other-

wise normal hearing, speech recognition, and a lack of obvious neurological lesions. Because amusics are also impaired at non-musical tasks that depend on pitch (Hyde and Peretz, 2004; Tillmann et al., 2009; Liu et al., 2010), their musical deficits are thought to reflect an underlying impairment in pitch perception and memory. For example, some amusics are unable to detect relatively large changes in pitch between isolated tones (e.g., >1 semitone, the pitch difference between adjacent keys on a piano; Hyde and Peretz, 2004).

Received July 16, 2015; revised Dec. 31, 2015; accepted Feb. 1, 2016.

Author contributions: S.V.N.-H., P.A., A.C., J.H.M., N.G.K., and B.T. designed research; S.V.N.-H. and P.A. performed research; S.V.N.-H. analyzed data; S.V.N.-H., P.A., A.C., J.H.M., N.G.K., and B.T. wrote the paper.

This work was supported by the Agence Nationale de la Recherche (ANR) of the French Ministry of Research (Grant ANR-11-BSH2-001-01 to B.T. and A.C.) and by the Centre National de la Recherche Scientifique. P.A. was funded by a doctoral fellowship from the Centre National de la Recherche Scientifique. This work was conducted in the framework of the LabEx CeLyA (Centre Lyonnais d'Acoustique; Grant ANR-10-LABX-0060) and of the LabEx Cortex (Construction, Function and Cognitive Function and Rehabilitation of the Cortex; Grant ANR-11-LABX-0042) of the Université de Lyon within the program Investissements d'avenir (Grant ANR-11-IDEX-0007) operated by the ANR. We thank Michael Cohen, Leyla Isik, and Caroline Robertson for helpful comments on an earlier draft of the manuscript. We also thank Ev Fedorenko for helpful discussions and for initiating the collaboration that led to this study.

The authors declare no competing financial interests.

Correspondence should be addressed to Sam Norman-Haignere, Department of Brain and Cognitive Sciences, MIT, 43 Vassar Street, 46-4141C, Cambridge, MA 02139. E-mail: snormanhaignere@gmail.com.

DOI:10.1523/JNEUROSCI.2705-15.2016

Copyright © 2016 the authors 0270-6474/16/362986-09\$15.00/0

The causes of amusia are actively debated. Psychophysical evidence suggests that amusics' pitch perception deficits cannot be explained by poor frequency resolution in the cochlea (Cousineau et al., 2015). Several neurophysiological studies suggest that white matter tracts linking auditory and frontal cortices are reduced in listeners with amusia (Hyde et al., 2006; Loui et al., 2009; Albouy et al., 2013; but see Chen et al., 2015), consistent with findings of abnormal functional and effective connectivity (Hyde et al., 2011; Albouy et al., 2013, 2015).

Less is known about whether impaired pitch coding in auditory cortex contributes to amusics' pitch deficits. An MEG study reported that evoked responses to each note of a melody (the N100m) were decreased in amplitude and temporally delayed in the auditory and frontal cortex of amusics, potentially related to findings of impaired effective connectivity between those same regions (Albouy et al., 2013). In contrast, EEG studies have reported seemingly normal deviance responses to pitch changes in the auditory cortex of amusics, even for pitch changes that are too small for amusics to detect (Moreau et al., 2013; Zendel et al., 2015). A prior fMRI study also reported that amusics, like normal listeners, show overall enhanced responses throughout auditory cortex to note sequences with pitch changes compared to note sequences with a fixed pitch (Hyde et al., 2011). However, given that amusics' deficits are specific to pitch, it seems plausible that their cortical impairments might also be specific to regions critical for pitch coding, rather than being distributed throughout auditory cortex. We thus sought to test whether the regions most implicated in pitch processing in normal-hearing listeners show abnormal responses in listeners with amusia.

Many prior studies have reported that normal-hearing listeners exhibit circumscribed cortical regions that respond preferentially to sounds with pitch (Patterson et al., 2002; Penagos et al., 2004; Norman-Haignere et al., 2013; Plack et al., 2014). These regions are found in a stereotyped anatomical location, partially overlapping the low-frequency field of primary auditory cortex. Although some prior studies have not reported clear selectivity for pitch in these anatomical regions (Hall and Plack, 2007; Barker et al., 2011, 2012), most of these studies have used pitch stimuli with "unresolved" harmonics, which produce a weak pitch percept relative to "resolved" harmonics (Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994), and a weak neural response in pitch-responsive cortical regions (Penagos et al., 2004; Norman-Haignere et al., 2013). Interestingly, amusic pitch deficits seem to reflect a selective impairment in the perception of resolved harmonics (Cousineau et al., 2015), which are the primary driver of pitch-responsive regions. Thus we sought to test whether pitch-responsive regions are detectable in the auditory cortex of individuals with congenital amusia using stimuli with resolved harmonics, and if so, whether they differ in pitch selectivity, extent, or anatomy from matched controls.

Materials and Methods

Participants. Twelve amusic subjects (ages 18–57, mean age = 37, five male and seven female) and 12 age- and education-matched controls (ages 18–53, mean age = 36, five male and seven female) participated in the experiment. There was a technical problem with stimulus presentation in one control subject, so the data from this subject and their matched amusic subject were excluded (all analyses and figures are based on the 11 remaining matched pairs). Amusics were identified with a standard musical aptitude test (the "Montreal Battery of Evaluation of Amusia" or MBEA; Fig. 1A). Head motion and voxel SNR were similar between the two groups, with no significant difference in either measure ($t_{(20)} < 0.2$, $p > 0.8$ for both). All subjects were right-handed and had audiometric thresholds < 30 dB HL for frequencies at or below 4 kHz. One control subject had mild hearing loss in her left ear above the fre-

quency range of the stimuli tested (45 and 50 dB HL at 6 and 8 kHz, respectively). None of the control subjects had formal musical training. Two amusic subjects had 1 year of training and 1 amusic subject had 5 years of training. The study was approved by the appropriate French local ethics committee on Human Research (CPP Sud-Est IV, number 12/100; 2012-A01209-34) and all participants gave written informed consent.

Behavioral pitch discrimination thresholds. We tested whether the amusic listeners in this study had higher-than-normal thresholds for detecting a change in pitch, as expected given prior work (Hyde and Peretz, 2004). This analysis was important because listeners were selected based on the results of a standard music aptitude test (the MBEA) rather than pitch discrimination performance. Thresholds for detecting a change in pitch for a 512 Hz pure tone were measured using a two-interval, forced-choice paradigm (for details, see Tillmann et al., 2009). Adaptive tracking was used to determine listeners' thresholds for detecting a pitch change (2-up, 1-down, targeting 71% performance).

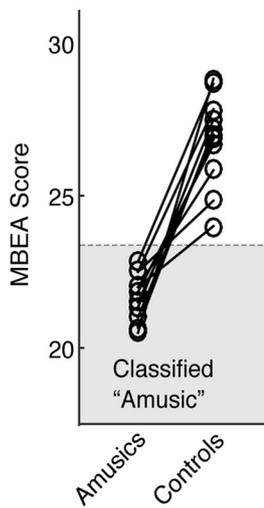
Stimuli for fMRI experiment. Stimuli were composed of either harmonic tone complexes or Gaussian noise (see Fig. 1B for a schematic of the design). The tone complexes contained harmonics 3–6 of their fundamental frequency (F0). We did not include the fundamental frequency or second harmonic in the stimulus because they are not needed to produce a robust pitch percept (Houtsma and Smurzynski, 1990) and because their inclusion produces an excitation pattern (the average cochlear response as a function of frequency) that more substantially differs from that of noise due to their wide spacing in the cochlea. Gaussian noise sounds were filtered to span the same frequency range as the harmonic tone complexes. Each stimulus lasted 2 s and included 6 "notes" that were varied in frequency to minimize adaptation (for details, see Norman-Haignere et al., 2013).

For each stimulus, the overall frequency range across all notes spanned either a low or high spectral region. We used two frequency ranges so that we could also test for tonotopic organization as a positive control in case amusic subjects showed weaker or absent pitch responses. Our analyses focused on characterizing pitch-responsive voxels by contrasting responses to harmonic tones and noise, combining across the two frequency ranges (frequency-selective responses reflecting tonotopy were evident in both groups, as expected). It is important to note that although pitch-responsive regions overlap the low-frequency area and respond preferentially to low-frequency sounds, pitch regions can nonetheless be identified using relatively high-frequency sounds, as they respond more to high-frequency harmonic tones than to high-frequency noise (Norman-Haignere et al., 2013). Thus, to assess pitch responses, we contrasted responses to harmonic tones and noise, and summed this contrast across both low- and high-frequency ranges to maximize statistical power: [low tones – low noise] + [high tones – high noise]. Similar results were obtained when we analyzed each frequency range separately (i.e., pitch-responsive regions in amusics and controls were of similar selectivity). The mean F0s for the low- and high-frequency harmonic notes were 166 and 666 Hz, respectively (yielding frequency ranges of the harmonics spanning 0.5–1 kHz and 2–4 kHz, respectively). We did not use noise to mask cochlear distortion products because for spectrally "resolvable" harmonics, like those tested here, distortion products have little effect on the response of pitch regions (Norman-Haignere and McDermott, 2016).

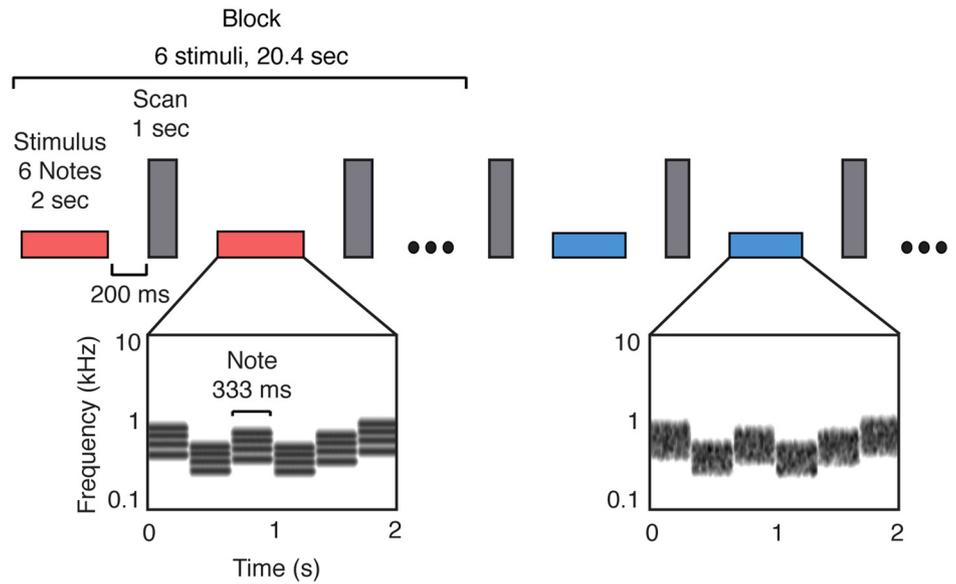
Procedure. Stimuli were presented in a sparse, blocked design, with 6 2 s stimuli from the same condition presented successively in each block (Fig. 1B). After each stimulus, a single scan was collected. To focus subjects' attention on the stimuli, participants performed a rhythm judgment task intended to be similarly difficult for amusics and controls: each stimulus had notes of either equal durations (333 ms) or irregular durations (183–583 ms), and subjects were instructed to indicate whether they heard a regular or irregular rhythm using a button press. Performance on the rhythm task was similar between amusics and controls, with no significant group difference ($t_{(20)} = 1.42$; $p = 0.17$).

Each subject completed a single run of the experiment, which included five blocks for each of the four conditions and five blocks of silence to provide a baseline with which to compare responses (each block lasted 20.4 s). Condition orders were pseudorandom and counterbalanced across subjects: for each subject, a set of condition orders was selected from a large set of ran-

A Musical Aptitude

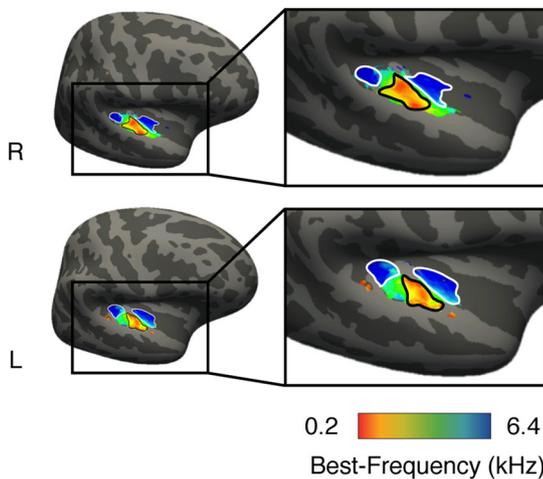


B Schematic of Experimental Design



C Tonotopy

Reproduced from Norman-Haignere et al. (2013)
Data from 12 normal-hearing subjects



D Pitch-Responsive Voxels

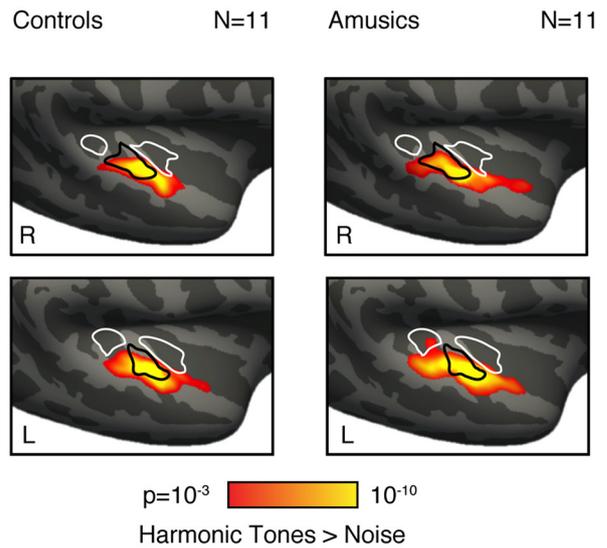


Figure 1. Schematic of experimental design and maps of pitch responses. **A**, Amusic subjects were identified using a standard musical aptitude test: the Montreal Battery of Evaluation of Amusia (MBEA). Each amusic subject was paired with a corresponding age- and education-matched control (matched pairs are indicated by connecting lines). **B**, Schematic of the experimental design. fMRI responses were measured to harmonic tones and Gaussian noise spanning the same frequency range. Contrasting responses to these stimuli in typical listeners reveals anatomically stereotyped “pitch-responsive” regions that overlap low- but not high-frequency tonotopic areas of primary auditory cortex (Norman-Haignere et al., 2013). Stimuli (denoted by horizontal bars) were presented in a block design, with six stimuli from the same condition presented successively in each block (red and blue indicate different conditions). Each stimulus (2 s) included several notes that varied in frequency to minimize adaptation. Cochleograms are shown for an example harmonic tone stimulus (red bar) and an example noise stimulus (blue bar). Cochleograms plot time–frequency decompositions, similar to a spectrogram, that summarize the cochlea’s response to sound. After each stimulus, a single scan was collected (vertical, gray bars). **C**, Group tonotopic map reproduced from Norman-Haignere et al. (2013) and used here for comparison with pitch-responsive voxels (**D**). Colors indicate the pure-tone frequency with the highest response in each voxel. Black and white outlines indicate regions of low- and high-frequency selectivity, respectively. **D**, Voxels with a significant response preference for harmonic tones compared with frequency-matched noise (cluster-corrected to $p < 0.001$). In both amusics and controls, we observed significant clusters of pitch-responsive voxels that partially overlapped the low-frequency area of primary auditory cortex.

domly generated orders (20,000) such that, on average, each condition was equally likely to occur at each point in the run and each condition was preceded equally often by every other condition in the experiment (as in our prior work; Norman-Haignere et al., 2013).

Presentation software (Neurobehavioral Systems) was used to present sounds in the scanner and record button responses. Sounds were pre-

sented at a fixed level (70 dB SPL) using MRI-compatible earphones (Nordic NeuroLab).

Data acquisition and preprocessing. All data were collected using a 3T Philips Achieva TX scanner with a 32-channel head coil (at the Lyon Neurological Hospital). The details of the scanning sequence were identical to that used in Norman-Haignere et al. (2013). Briefly, each func-

tional volume (e.g., a single 3D image) comprised 15 slices covering most of the superior temporal cortex and oriented parallel to the superior temporal plane (slices were 4 mm thick with a 2.1×2.1 mm in-plane resolution). Volumes were acquired every 3.4 s. Each acquisition lasted 1 s and stimuli were presented in the 2.4 s gap of silence between acquisitions (Fig. 1B).

Functional volumes were motion corrected and aligned to the anatomical volume from each subject. The aligned volumes were resampled to the high-density surface mesh computed by FreeSurfer for each individual subject; and these individual-subject meshes were aligned to the mesh of a standardized template brain (the MNI305 FsAverage brain). After alignment, the mesh data were smoothed using a relatively small kernel (3 mm FWHM) and interpolated to a 1.5×1.5 mm grid using a flattened representation of the surface mesh. We used a slightly larger smoothing kernel to compute the group-averaged, whole-brain maps described below (5 mm FWHM) to account for the local variability of cortical responses across subjects.

Regression analyses. Each voxel was fit with a general linear model (GLM), with one regressor per stimulus condition. The regressors for each stimulus condition were computed in the standard way, using an estimate of the hemodynamic response function (HRF). This HRF estimate was calculated from the data using a finite-impulse response (FIR) model, rather than assuming a fixed parametric form (see “Estimating the hemodynamic response function” below). To model sources of noise, we included the following nuisance regressors: a linear-trend regressor (to account for signal drift) and the first 10 principal components from voxel responses in white matter (to account for sources of noise with high variance across voxels). This approach is similar to standard denoising techniques (Kay et al., 2013).

Estimating the hemodynamic response function. A simple way to improve GLM fits is to estimate the HRF using a FIR model (Glover, 1999). Each time point in the HRF was modeled with a separate “candlestick” regressor, with a 1 for all scans that occurred a fixed time delay after the onset of a stimulus block (regardless of stimulus type/condition) and a 0 for all other scans. These candlestick regressors were fit to each voxel’s response using ordinary least squares. The weights for each regressor, which collectively provide an estimate of each voxel’s HRF, were then averaged across voxels and subjects. We averaged responses across the 10% of voxels in the superior temporal plane (the anatomical region most responsive to sound) of each subject that were best explained by the candlestick regressors (the estimated HRF was robust to the exact number of voxels selected; e.g., selecting the top 50% of voxels yielded similar results). This analysis provided an estimate of the average HRF to a stimulus block in our experiment across all conditions and subjects. Regressors for each condition and each subject were computed from this HRF and fit to the voxel responses.

Whole-brain contrast maps. We calculated maps showing voxels with a significant response preference for sounds with pitch (harmonic tones > noise; Fig. 1D). Each voxel’s response time course was fit with the four stimulus regressors and 11 nuisance regressors described above. The weights for the tone and noise regressors were subtracted and then summed across the two frequency ranges (i.e., [low tones – low noise] + [high tones – high noise]). This difference score for each voxel and subject was converted to a z-statistic (using ordinary least-squares equations) and then averaged across subjects. To calculate group significance maps, we used a permutation test analogous to a standard, cluster-corrected significance test based on the familywise error rate (Nichols and Holmes, 2002). We used a permutation test because it does not require any assumptions about the distribution of voxel responses.

The permutation test was implemented using a standard procedure (Nichols and Holmes, 2002). The key idea of a permutation test is to create a null distribution by permuting the predicted correspondence between sets of variables. Here, we permuted the order of blocks for different stimulus conditions (e.g., A, B, C, D → D, C, A, B, where letters represent blocks from different conditions) so that there was no regular mapping between the actual and assumed onset times for different stimulus conditions. In each subject, we re-ran the regression analysis using stimulus regressors computed from each of 10,000 different permuted block orders (each subject had a different set of permuted orders), yield-

ing 10,000 z-statistic maps for the harmonic tone > noise contrast. We then randomly selected a single, permuted order from each of the 11 subjects in each group, and averaged the corresponding z-statistic maps across the 11 subjects. Using this procedure, we calculated 10,000 permuted z-statistic maps and compared this null distribution with the observed z-statistic maps to assess significance. Specifically, for each voxel, we fit the 10,000 z-statistics from the null distribution with a Gaussian and calculated the likelihood of obtaining the observed z-statistic (based on the unpermuted condition orders) given a sample from this Gaussian. The Gaussian fits made it possible to estimate small *p* values (e.g., $p = 10^{-10}$) that would be impossible to directly approximate by counting the fraction of samples from the null that exceeded the observed statistic (and for voxels with larger *p* values, the Gaussian fits yielded very similar *p* values to those computed via counting).

The approach just described was used to calculate voxelwise significance values (plotted in Fig. 1D). To correct for multiple comparisons, we used a simple variant of cluster-correction suited for the permutation test (Nichols and Holmes, 2002). For each set of permuted condition orders, we computed a map of voxelwise significance values using the permutation analysis just described. We then thresholded this uncorrected voxelwise significance map ($p < 0.001$) and recorded the size of the largest contiguous voxel cluster that exceeded this threshold (within a large constraint region, spanning the entire superior temporal plane and gyrus). Using this approach, we built up a null distribution for cluster sizes across the 10,000 permutations. To evaluate significance, we counted the fraction of times that the cluster sizes for this null distribution exceeded that for each observed cluster (based on unpermuted orders and the same $p < 0.001$ voxelwise threshold).

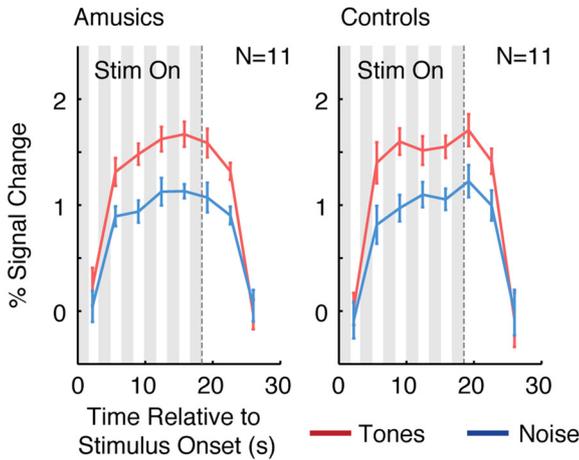
To compare the two groups, we contrasted their pitch responses (i.e., [amusia tones – amusia noise] vs [control tones – control noise]). We used the identical permutation analysis to assess significance except that we permuted the assignment of subjects to groups instead of the assignment of conditions to blocks. We also used a more liberal voxelwise significance threshold ($p < 0.05$) to maximize the chance of finding a cluster with a significant group difference (there were no voxels that passed the more stringent $p < 0.001$ threshold). Cluster correction was again used to assess whether the voxels that passed the liberal $p < 0.05$ threshold were significant when correcting for multiple comparisons.

Functional ROI analyses. We used a standard “localize-and-test” ROI analysis to measure both the extent and selectivity of pitch responses in amusics and controls. There were two parts to the analysis. First, we localized voxels in auditory cortex that showed the most consistent response preference for harmonic tones compared with noise (smallest *p* value for the harmonic tones > noise contrast) using a subset of our data. We then measured the response of the selected voxels to each stimulus condition in a disjoint subset of the data, to test whether they replicated their response selectivity for harmonic tones. This analysis was implemented by dividing each run into “subruns,” each of which included data from five blocks: one stimulus block from each of the four conditions and one block of silence. To calculate ROIs, we used the same “leave-one-out” analysis described in Norman-Haignere et al. (2013), but applied the analysis to subruns instead runs, in all cases using non-overlapping scans to select and measure the response of pitch voxels.

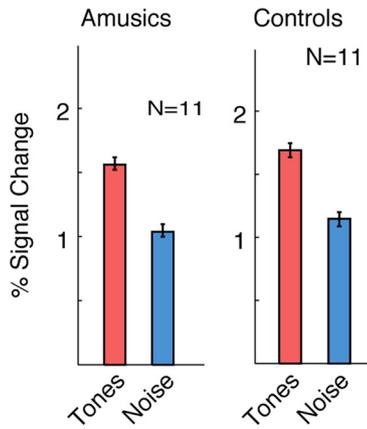
The analysis has one free parameter, the number of voxels selected, which determines the size of the ROI. In practice, selecting a smaller number of voxels that are more consistently responsive to pitch leads to more selective responses when measured in independent data. We thus measured responses from ROIs with a fixed number of voxels from each subject (the 10% of voxels in the superior temporal plane that were most pitch responsive, as in Norman-Haignere et al., 2013; Fig. 2) and also systematically probed the effect of ROI size on pitch selectivity (Fig. 3). This approach provides a simple way to jointly characterize both the extent and selectivity of pitch responses, which is necessary given that size and selectivity trade-off. To compare the selectivity of amusic and control ROIs quantitatively, we again used a permutation test that randomly permuted the assignment of subjects to groups (using the same approach described above).

Because overall response magnitudes can vary substantially across subjects for reasons unrelated to neural activity (e.g., vascularization), we com-

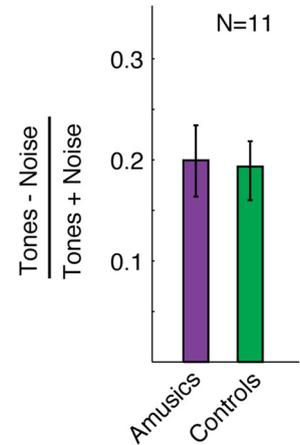
A Response Timecourse of Pitch Voxels



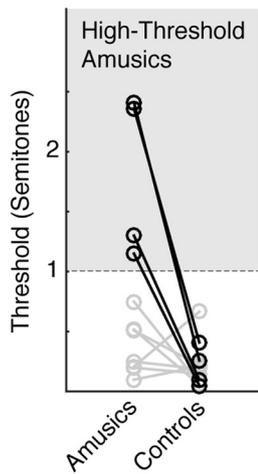
B Time-Averaged Response of Pitch Voxels



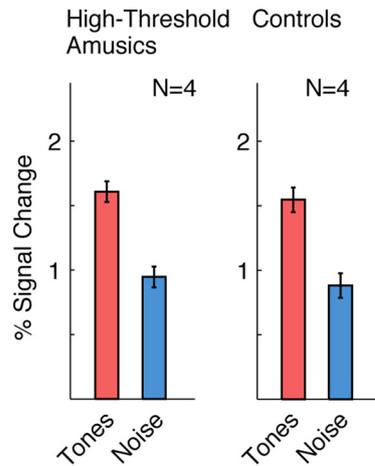
C Selectivity of Pitch Voxels



D Pitch Discrimination Thresholds



E Response of Pitch Voxels for Amusics with High Discrimination Thresholds



F Selectivity of Pitch Voxels for Amusics with High Thresholds

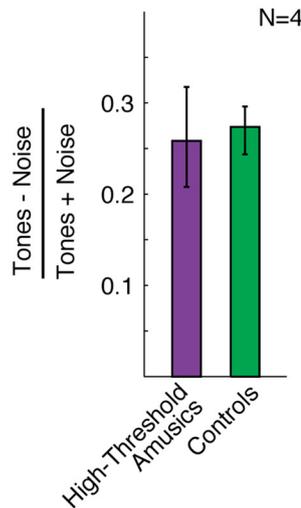


Figure 2. Pitch selectivity measured using functional ROIs. **A**, Response time course of pitch-responsive voxels to harmonic tones and noise, measured using an ROI analysis (see in “Functional ROI analyses” in the Materials and Methods). Gray regions indicate times when the stimulus was being played. The response to the stimulus is delayed because the BOLD response builds up slowly in response to neural activity. **B**, Time-averaged response to each condition (calculated using an HRF). **C**, Selectivity of the time-averaged response for harmonic tones compared with noise, measured by dividing the difference between responses to the two types of conditions by their sum. **D**, Pitch discrimination thresholds. Amusics’ thresholds were higher on average than controls, but the amusic population was heterogeneous, consistent with prior reports (Tillmann et al., 2009; Liu et al., 2010). Four amusic subjects (bolded circles) had particularly high thresholds (> 1 semitone) and were analyzed separately. **E**, Time-averaged response of pitch voxels for high-threshold amusics and their matched controls (same format as **B**). **F**, Selectivity of pitch-responsive voxels in high-threshold amusics and their matched controls (same format as **C**). Error bars in **A**, **B**, and **E** represent 1 SE of the mean difference between responses to harmonic tones and noise across subjects (computed via bootstrap); error bars in **C** and **F** represent 1 SE of the plotted selectivity measure across subjects (computed via bootstrap).

puted error bars and statistics on response differences between conditions. Figures 2A, 2B, 2E, and 3A plot 1 SE of the mean difference in response to harmonic tones and noise across subjects from each group (i.e., “within-subject” SEs; Loftus and Masson, 1994). As a result, these error bars should not be used to compare absolute response magnitudes between the two groups (and the two groups are thus plotted on different axes). Figures 2, C and F, and 3B plot response differences to harmonic tones and noise (normalized by their sum) and it is therefore appropriate to compare this index directly across groups (and the two groups are thus plotted on the same axis). In all cases, we used bootstrapping to calculate SEs.

Anatomical ROI analyses. We used two sets of anatomical ROIs to probe the distribution of pitch-responsive voxels (Fig. 4; Norman-Haignere et al., 2013). One set reflects standard anatomical subdivisions of auditory cortex based on macro-anatomical landmarks (Morosan et al., 2001; Fig. 4A, B). The second set are custom ROIs designed to run along the posterior-to-anterior axis of the superior temporal plane and to each include an equal number of sound-responsive voxels (Fig. 4C, D). The most anterior ROI is slightly larger in extent than the others because the anatomical region that it covers has fewer sound-responsive voxels. For each ROI, we measured the fraction of sound-responsive voxels that

A Response of Pitch Voxels for ROIs of Varying Size

B Pitch Selectivity vs ROI Size

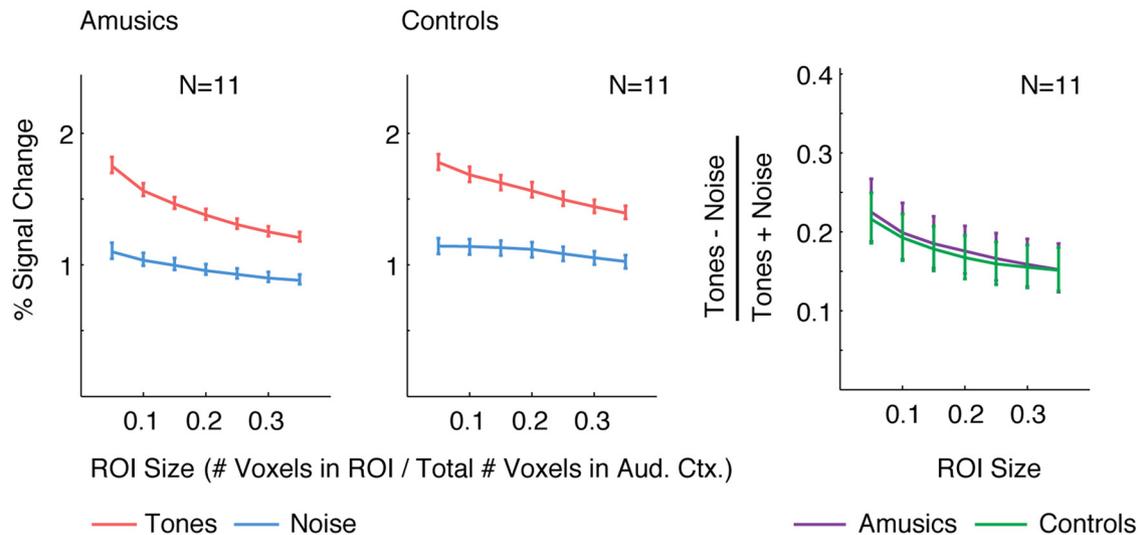


Figure 3. Extent of pitch selectivity measured using functional ROIs of varying size. **A**, Response of ROIs of varying size to harmonic tones and noise. Each ROI included the top N% of voxels in auditory cortex with the most significant response preference for harmonic tones compared with noise (selected using independent data from that used to measure the ROI's response). Error bars represent 1 SE of the mean difference between responses to harmonic tones and noise across subjects (computed via bootstrap). **B**, Selectivity of each ROI for harmonic tones compared with noise. Error bars represent 1 SE of the plotted selectivity measure across subjects (computed via bootstrap).

also had a pitch response, as a measure of the density of pitch-responsive voxels in each region. Pitch-responsive voxels were again defined as the 10% of voxels in auditory cortex with the most significant pitch response (harmonic tone > noise contrast). Sound-responsive voxels were defined by contrasting responses to both harmonic tones and noise with silence ($p < 0.05$). Measuring the density of pitch responses in this way controls for differences in the overall size of each ROI (Norman-Haignere et al., 2013). To statistically compare the density of pitch responses between groups, we measured the squared deviation between the density of pitch responses in amusics and controls, averaged across the ROIs from each set. We then compared this average deviation score with that for a null distribution calculated by permuting the assignments of subjects to groups. Significance was determined by counting the fraction of times the deviation scores from the null distribution exceeded the observed value. SEs were calculated using bootstrapping across subjects.

Results

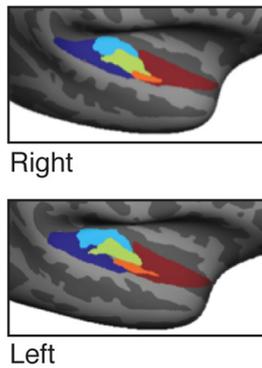
Figures 1C and 1D show voxels in amusics and controls that responded significantly more to harmonic tones compared with frequency-matched noise, superimposed on contours from a tonotopic map. We used the widely reported high-low-high tonotopic gradient of primary auditory cortex (Formisano et al., 2003; Humphries et al., 2010; Da Costa et al., 2013; Norman-Haignere et al., 2013) as a functional landmark with which to compare responses to pitch. As expected, we observed bilateral clusters of voxels in control subjects that responded significantly more to harmonic tones compared with noise (cluster-corrected $p < 0.001$, via a permutation test; see Materials and Methods) and that partially overlapped the low-frequency field of primary auditory cortex (Fig. 1C,D). These results in normal listeners replicate prior reports (Patterson et al., 2002; Norman-Haignere et al., 2013). Our key finding is that amusic listeners also exhibited highly significant pitch-responsive voxels in a similar location (cluster-corrected $p < 0.001$). The extent of pitch responses in amusics and controls was qualitatively similar. Although there

were some voxels that passed our statistical cutoff ($p < 0.001$) in one group but not the other, these group differences plausibly reflect noise because there was no significant group difference after cluster correction ($p = 0.41$ via permutation test).

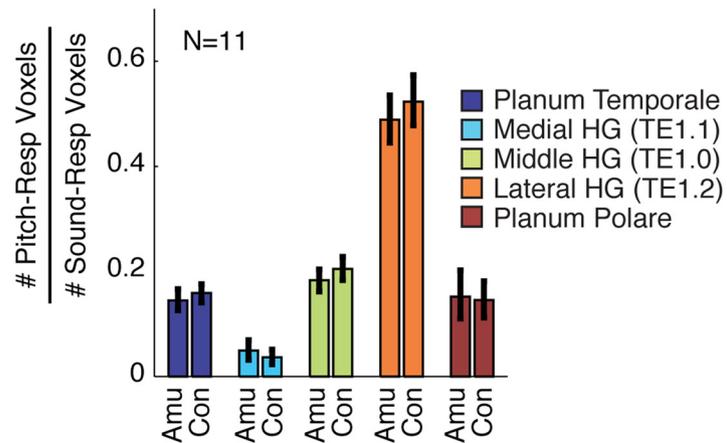
The selectivity of pitch-responsive voxels was quantified with an ROI analysis (Fig. 2), using independent data to define and measure each ROI's response (see "Functional ROI analyses" in Materials and Methods). Pitch-responsive ROIs in amusics and controls exhibited a similar response preference for harmonic tones compared with noise that was present throughout the duration of each ROI's response (Fig. 2A,B). Moreover, every amusic and every control subject tested showed a greater response (time averaged over the block) to harmonic tones compared with noise, leading to a highly significant response preference for harmonic tones in each group. We quantified the selectivity of this response preference using a standard selectivity index applied to each ROI's time-averaged response ($[\text{tones} - \text{noise}]/[\text{tones} + \text{noise}]$; Fig. 2C). This analysis revealed that amusics' ROIs were just as selective for pitch as controls' ROIs, with no significant group difference ($p = 0.88$, permutation test). There was also no significant difference between amusics and controls for ROIs computed separately for each hemisphere ($p > 0.8$ for both left and right hemisphere ROIs, permutation test).

We next investigated whether robust pitch-responsive voxels could be detected in a subset of amusic subjects who were particularly impaired at discriminating pitch (defined as having discrimination thresholds >1 semitone; Fig. 2D). Even these amusic subjects exhibited pitch-responsive voxels that were just as selective as controls (Fig. 2E,F; $p = 0.8$, permutation test). This analysis was important because some amusics had thresholds in the range of the control subjects (consistent with prior work; Tillmann et al., 2009; Liu et al., 2010) even though, on average, pitch-discrimination thresholds were

A Standard Anatomical ROIs

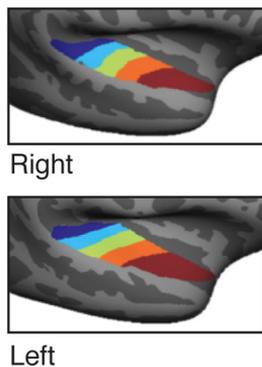


B Density of Pitch-Responsive Voxels in Each ROI



C Posterior-to-Anterior Anatomical ROIs

ROIs with Equally Many Sound-Responsive Voxels



D Density of Pitch-Responsive Voxels in Each ROI

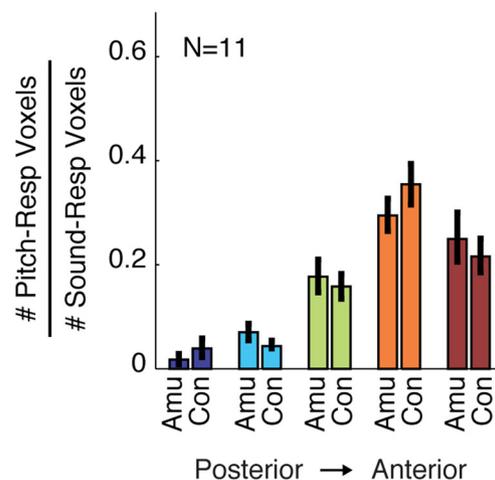


Figure 4. Anatomical distribution of pitch-responsive voxels. **A**, Five anatomical ROIs subdividing auditory cortex into standard regions (Morosan et al., 2001). Three ROIs subdivided Heschl's gyrus (HG) based on cytoarchitecture. Prior studies have reported a high density of pitch-responsive voxels in lateral HG (Patterson et al., 2002). The planum temporale and planum polare demarcate regions posterior and anterior to HG, respectively. **B**, Density of pitch-responsive voxels in each anatomical ROI, measured by the fraction of sound-responsive voxels (harmonic tones + noise > silence) that also exhibited a pitch response (harmonic tones > noise). **C**, Anatomical ROIs designed to run along the posterior-to-anterior axis of the superior temporal plane and to each include an equal number of sound-responsive voxels (Norman-Haignere et al., 2013). Prior work using these ROIs has shown a high density of pitch-responsive voxels in anterior regions of auditory cortex (Norman-Haignere et al., 2013). **D**, Density of pitch responses in each posterior-to-anterior ROI. Error bars in **B** and **D** represent 1 SE of the plotted density measure across subjects (computed via bootstrap).

higher in the amusic group compared with the control group ($t_{(20)} = 2.60, p < 0.05$).

We next analyzed the extent of pitch responses by measuring pitch selectivity for ROIs of varying size (Fig. 3). Each ROI included the top N% of voxels in auditory cortex with the most consistent response preference for harmonic tones compared with noise (selected using independent data from that used to measure the ROI's response, see "Functional ROI analyses" in the Materials and Methods). As expected, smaller ROIs exhibited more selective responses because they were composed of voxels with a more consistent response to pitch. However, the selectivity of amusic and control ROIs was closely matched for all sizes tested ($p > 0.8$ for all ROI sizes tested, permutation test).

Finally, we assessed the anatomical distribution of pitch responses using two sets of anatomical ROIs (Fig. 4). We quantified the density of pitch responses in each ROI as the proportion of sound-responsive voxels (harmonic tones + noise > silence) that also exhibited a pitch response (harmonic tones > noise; see "Anatomical ROI analyses" in the Materials and Methods). For all ROIs, the density of pitch responses in amusic and controls was similar ($p > 0.4$ for both sets of ROIs, permutation test). Both amusic and controls exhibited a high density of pitch responses in lateral HG (Fig. 4A,B) and an overall bias for pitch-responsive voxels to be located in anterior regions of auditory cortex (Fig. 4C,D), replicating prior work in normal listeners (Patterson et al., 2002; Norman-Haignere et al., 2013).

Discussion

Our results demonstrate pitch-responsive regions in amusic individuals that are comparable in extent, selectivity, and anatomical location to those of controls. Highly significant clusters of pitch-responsive voxels were observed in amusics, overlapping the low-frequency area of lateral Heschl's gyrus as in typical listeners (Patterson et al., 2002; Penagos et al., 2004; Norman-Haignere et al., 2013; Plack et al., 2014). Pitch-responsive voxels in amusics were just as selective for pitch as those in controls, even in amusic subjects who were unable to discriminate large changes in pitch that would be obvious to a normal listener (i.e., pitch changes >1 semitone).

These findings raise the question of why listeners with seemingly normal pitch-responsive cortical regions have such difficulty perceiving pitch. One possibility, suggested by prior work, is that listeners may be unable to consciously access pitch information that is encoded in their auditory cortex. For example, prior studies have reported that pitch changes too small for amusic listeners to overtly detect nonetheless produce detectable neural responses in their auditory cortex, as measured with both EEG (Moreau et al., 2013; Zendel et al., 2015) and fMRI (Hyde et al., 2011). Consistent with these neural measurements, implicit behavioral measures have revealed sensitivity to musical pitch structure that is difficult to detect with explicit measures (Omigie et al., 2012; Tillmann et al., 2012). For example, amusic listeners exhibit a normal priming effect of note expectancy (based on the musical pitch context) on performance for an unrelated timbre discrimination task (faster reaction times for expected notes), but are substantially impaired relative to controls at judging which notes are expected versus unexpected (on a 1–7 scale) when asked explicitly (Omigie et al., 2012). In addition, amusics' pitch deficits are exacerbated when pitch information must be retained in short-term memory across a delay (Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010). These deficits in perceptual awareness and memory for pitch have been argued to stem from impaired connectivity between auditory cortex and frontal regions critical to explicit task performance and short-term memory (Loui et al., 2009; Hyde et al., 2011; Albouy et al., 2013; but see Chen et al., 2015). Thus, amusic deficits in pitch and music perception could be caused by impairments in the fronto-temporal pathway (Hyde et al., 2011; Albouy et al., 2013, 2015), even if pitch information is faithfully encoded within auditory cortex.

Another possibility is that neurons within pitch-responsive regions could exhibit tuning differences in amusics not evident in the region's overall response. For example, the tuning of cortical neurons for different fundamental frequencies could be coarser in amusics, perhaps leading to impaired pitch discrimination. It is also possible that relevant information might be encoded in the temporal characteristics of the neural response in pitch regions and other areas, which would be difficult to detect with fMRI. For example, Albouy et al. (2013) provided evidence that cortical MEG responses to tones are delayed in amusics relative to controls and Lehmann et al. (2015) reported that, even in the auditory brainstem, responses to voiced syllables are temporally delayed in amusics.

A third possibility is that amusic deficits might reflect impairments in regions outside of pitch-responsive cortex. For example, several recent studies have reported music-selective responses in human auditory cortex (Leaver and Rauschecker, 2010; Angulo-Perkins et al., 2014; Norman-Haignere et al., 2015). These responses are located outside of the pitch-responsive regions

studied here, and cannot be explained as a response to standard acoustic features such as frequency or spectrotemporal modulation. Given that some amusics have deficits in music perception, but not pitch discrimination (Tillmann et al., 2009; Liu et al., 2010), it is plausible that impairments to music-selective circuitry could underlie some amusic deficits. In addition, there is evidence that subregions of the intraparietal sulcus are important for the active maintenance of musical information in working memory (Foster et al., 2013) and thus might play a role in amusics' short-term memory deficits (Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010; Albouy et al., 2013). Prior studies have also reported abnormal responses to tones in the right frontal cortex of amusics (Hyde et al., 2011; Albouy et al., 2013), which could be related to reported impairments in connectivity between frontal cortex and auditory areas (Loui et al., 2009; Hyde et al., 2011; Albouy et al., 2013).

Critically, the ability to identify pitch-responsive areas in amusics (along with other functionally specific regions) will make it possible to test these different possibilities in future work. Methods such as fMRI adaptation and multivoxel pattern analysis could be used to detect differences in the tuning of neurons within a functionally defined region. It will also be possible to test whether reductions in connectivity between auditory cortex and frontal cortex are greater for regions preferentially responsive to pitch, as might be expected given the specificity of amusics' behavioral deficits for pitch perception and memory. For example, the connectivity of pitch-responsive regions with frontal cortex could be measured and compared with that of other functionally specific regions, like those selective for voice and speech sounds (Belin et al., 2000). There are also emerging techniques that make it possible to measure the temporal response of functionally defined regions by combining fMRI with methods that have improved temporal resolution, such as MEG (Baldauf and Desimone, 2014).

Our findings on amusia bear some intriguing similarities to prior work on other neurodevelopmental disorders. For example, children with dyslexia have great difficulty with tasks that require overt awareness, manipulation, and memory for phonological structures (e.g., remembering a sequence of phonemes, saying "game" without "g"; Ramus and Szenkovits, 2008; Norton et al., 2015), but responses to phonemes in speech-selective cortical areas have been reported to be largely intact (Boets et al., 2013). Similarly, robust face-selective responses have been observed in subjects with congenital face-recognition deficits ("prosopagnosia"; Hasson et al., 2003; Zhang et al., 2015), although small differences in selectivity and size may exist (Furl et al., 2011), especially in more anterior regions (Avidan et al., 2014). Moreover, several studies have suggested that deficits in white matter connectivity may play a key role in both dyslexia (Boets et al., 2013) and prosopagnosia (Thomas et al., 2009; Gomez et al., 2015), similar to prior work on amusia (Loui et al., 2009; Hyde et al., 2011; Albouy et al., 2013). Therefore, one speculative possibility is that congenital deficits across different domains may be caused by similar types of neurological impairments, so by better understanding amusia, we may gain deeper insights into the causes of neurodevelopmental disorders more broadly.

References

- Albouy P, Mattout J, Bouet R, Maby E, Sanchez G, Aguera PE, Daligault S, Delpuech C, Bertrand O, Caclin A, Tillmann B (2013) Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex. *Brain* 136:1639–1661. [CrossRef Medline](#)
- Albouy P, Mattout J, Sanchez G, Tillmann B, Caclin A (2015) Altered retrieval of melodic information in congenital amusia: insights from

- dynamic causal modeling of MEG data. *Front Hum Neurosci* 9:20. [CrossRef Medline](#)
- Angulo-Perkins A, Aubé W, Peretz I, Barrios FA, Armony JL, Concha L (2014) Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. *Cortex* 59:126–137. [CrossRef Medline](#)
- Avidan G, Tanzer M, Hadj-Bouziane F, Liu N, Ungerleider LG, Behrmann M (2014) Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. *Cereb Cortex* 24:1565–1578. [CrossRef Medline](#)
- Baldauf D, Desimone R (2014) Neural mechanisms of object-based attention. *Science* 344:424–427. [CrossRef Medline](#)
- Barker D, Plack CJ, Hall DA (2011) Human auditory cortical responses to pitch and to pitch strength. *Neuroreport* 22:111–115. [CrossRef Medline](#)
- Barker D, Plack CJ, Hall DA (2012) Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. *Cereb Cortex* 22:745–753. [CrossRef Medline](#)
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403:309–312. [CrossRef Medline](#)
- Boets B, Op de Beeck HP, Vandermosten M, Scott SK, Gillebert CR, Mantini D, Bulthé J, Sunaert S, Wouters J, Ghesquière P (2013) Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342:1251–1254. [CrossRef Medline](#)
- Chen JL, Kumar S, Williamson VJ, Scholz J, Griffiths TD, Stewart L (2015) Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Front Psychol* 6:9. [CrossRef Medline](#)
- Cousineau M, Oxenham AJ, Peretz I (2015) Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia* 66:293–301. [CrossRef Medline](#)
- Da Costa S, van der Zwaag W, Miller LM, Clarke S, Saenz M (2013) Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. *J Neurosci* 33:1858–1863. [CrossRef Medline](#)
- Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R (2003) Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* 40:859–869. [CrossRef Medline](#)
- Foster NE, Halpern AR, Zatorre RJ (2013) Common parietal activation in musical mental transformations across pitch and time. *Neuroimage* 75:27–35. [CrossRef Medline](#)
- Furl N, Garrido L, Dolan RJ, Driver J, Duchaine B (2011) Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *J Cogn Neurosci* 23:1723–1740. [CrossRef Medline](#)
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9:416–429. [CrossRef Medline](#)
- Gomez J, Pestilli F, Witthoft N, Golarai G, Liberman A, Poltoratski S, Yoon J, Grill-Spector K (2015) Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. *Neuron* 85:216–227. [CrossRef Medline](#)
- Gosselin N, Jolicœur P, Peretz I (2009) Impaired memory for pitch in congenital amusia. *Ann N Y Acad Sci* 1169:270–272. [CrossRef Medline](#)
- Hall DA, Plack CJ (2007) The human “pitch center” responds differently to iterated noise and Huggins pitch. *Neuroreport* 18:323–327. [CrossRef Medline](#)
- Hasson U, Avidan G, Deouell LY, Bentin S, Malach R (2003) Face-selective activation in a congenital prosopagnosic subject. *J Cogn Neurosci* 15:419–431. [CrossRef Medline](#)
- Houtsma AJM, Smurzynski J (1990) Pitch identification and discrimination for complex tones with many harmonics. *J Acoust Soc Am* 87:304–310. [CrossRef Medline](#)
- Humphries C, Liebenthal E, Binder JR (2010) Tonotopic organization of human auditory cortex. *Neuroimage* 50:1202–1211. [CrossRef Medline](#)
- Hyde KL, Peretz I (2004) Brains that are out of tune but in time. *Psychol Sci* 15:356–360. [CrossRef Medline](#)
- Hyde KL, Zatorre RJ, Griffiths TD, Lerch JP, Peretz I (2006) Morphometry of the amusic brain: a two-site study. *Brain* 129:2562–2570. [CrossRef Medline](#)
- Hyde KL, Zatorre RJ, Peretz I (2011) Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cereb Cortex* 21:292–299. [CrossRef Medline](#)
- Kay KN, Rokem A, Winawer J, Dougherty RF, Wandell BA (2013) GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Front Neurosci* 7:247. [CrossRef Medline](#)
- Leaver AM, Rauschecker JP (2010) Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J Neurosci* 30:7604–7612. [CrossRef Medline](#)
- Lehmann A, Skoe E, Moreau P, Peretz I, Kraus N (2015) Impairments in musical abilities reflected in the auditory brainstem: evidence from congenital amusia. *Eur J Neurosci* 42:1644–1650. [CrossRef Medline](#)
- Liu F, Patel AD, Fourcin A, Stewart L (2010) Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain* 133:1682–1693. [CrossRef Medline](#)
- Loftus GR, Masson ME (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1:476–490. [CrossRef Medline](#)
- Loui P, Alsop D, Schlaug G (2009) Tone deafness: a new disconnection syndrome? *J Neurosci* 29:10215–10220. [CrossRef Medline](#)
- Moreau P, Jolicœur P, Peretz I (2013) Pitch discrimination without awareness in congenital amusia: evidence from event-related potentials. *Brain Cogn* 81:337–344. [CrossRef Medline](#)
- Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13:684–701. [Medline](#)
- Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* 15:1–25. [CrossRef Medline](#)
- Norman-Haignere S, Kanwisher N, McDermott JH (2013) Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. *J Neurosci* 33:19451–19469. [CrossRef Medline](#)
- Norman-Haignere S, Kanwisher NG, McDermott JH (2015) Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron* 88:1281–1296. [CrossRef Medline](#)
- Norman-Haignere S, McDermott JH (2016) Distortion products in auditory fMRI research: measurements and solutions. *Neuroimage*. Advance online publication. [CrossRef Medline](#)
- Norton ES, Beach SD, Gabrieli JD (2015) Neurobiology of dyslexia. *Curr Opin Neurobiol* 30:73–78. [CrossRef Medline](#)
- Omigie D, Pearce MT, Stewart L (2012) Tracking of pitch probabilities in congenital amusia. *Neuropsychologia* 50:1483–1493. [CrossRef Medline](#)
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36:767–776. [CrossRef Medline](#)
- Penagos H, Melcher JR, Oxenham AJ (2004) A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci* 24:6810–6815. [CrossRef Medline](#)
- Peretz I, Ayotte J, Zatorre RJ, Mehler J, Ahad P, Penhune VB, Jutras B (2002) Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* 33:185–191. [CrossRef Medline](#)
- Plack CJ, Barker D, Hall DA (2014) Pitch coding and pitch processing in the human brain. *Hear Res* 307:53–64. [CrossRef Medline](#)
- Ramus F, Szenkovits G (2008) What phonological deficit? *Q J Exp Psychol (Hove)* 61:129–141. [CrossRef Medline](#)
- Shackleton TM, Carlyon RP (1994) The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J Acoust Soc Am* 95:3529–3540. [CrossRef Medline](#)
- Thomas C, Avidan G, Humphreys K, Jung KJ, Gao F, Behrmann M (2009) Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat Neurosci* 12:29–31. [Medline](#)
- Tillmann B, Schulze K, Foxton JM (2009) Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain Cogn* 71:259–264. [CrossRef Medline](#)
- Tillmann B, Gosselin N, Bigand E, Peretz I (2012) Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex* 48:1073–1078. [CrossRef Medline](#)
- Williamson VJ, McDonald C, Deutsch D, Griffiths TD, Stewart L (2010) Faster decline of pitch memory over time in congenital amusia. *Adv Cogn Psychol* 6:15–22. [CrossRef Medline](#)
- Zendel BR, Lagrois MÉ, Robitaille N, Peretz I (2015) Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *J Neurosci* 35:3815–3824. [CrossRef Medline](#)
- Zhang J, Liu J, Xu Y (2015) Neural decoding reveals impaired face configural processing in the right fusiform face area of individuals with developmental prosopagnosia. *J Neurosci* 35:1539–1548. [CrossRef Medline](#)