Pitch-Responsive Cortical Regions in Congenital Amusia

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/DOI:10.1523/JNEUROSCI.2705-15.2016">http://dx.doi.org/DOI:10.1523/JNEUROSCI.2705-15.2016</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Biomed Central Ltd.</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sat Dec 29 12:59:16 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/108659">http://hdl.handle.net/1721.1/108659</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution 4.0 International License</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a> Adam Science Foundation (2007).  This license allows for Creative Commons Attribution 4.0 International License.</td>
</tr>
</tbody>
</table>
Pitch-Responsive Cortical Regions in Congenital Amusia

Sam V. Norman-Haignere, Philippe Albouy, Anne Caclin, Josh H. McDermott, Nancy G. Kanwisher, and Barbara Tillmann

Department of Brain and Cognitive Sciences and McGovern Institute for Brain Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. Lyon Neuroscience Research Centre, Centre National de la Recherche Scientifique UMR5292, INSERM U1028, Lyon 1 University, F-69000 Lyon, France, and Montreal 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 amusiac 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 amusics 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, amusics 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 amusiac subjects will make it possible to ask more precise questions about their role in amusiac 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 otherwise 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.


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 Federenko 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.


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; Slackleton and Carlyon, 1994), and a weak neural response in pitch-responsive cortical regions (Penagos et al., 2004; Norman-Haignere et al., 2013). Interestingly, amusics' 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(38) < 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 frequency range of the stimuli tested (45 and 30 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 (CPR Sud-Est IV, number 12/100; 2012-A01209-34) and all participants gave written informed consent.

Behavioral pitch discrimination thresholds. We tested whether the amusics 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 voxels were also similar with respect to “selectivity”). The mean Pfs 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(119) = 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-
randomly 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 presented 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 $2.1 \times 2.1$ 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 MNI152 FsAverage brain). After alignment, the mesh data were smoothed using a relatively small kernel (3 mm FWHM) and interpolated to a $1.5 \times 1.5$ 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 variate 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 “candelstick” 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 regressors for all other scans. These candlestick regressors were fit to each voxel’s response time course to compute the group-averaged, whole-brain maps described below (5 mm FWHM) to account for the local variability of cortical responses across subjects.

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 ($p < 0.005$) 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 described above; 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. For each subject, they used 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 amusics and control ROIs, we again used a permutation test that randomly permuted the assignment of subjects to groups (as described above) and selected a smaller number of voxels ($n < 0.005$ voxelwise threshold), whether they replicated their response selectivity for harmonic tones. Because overall response magnitudes can vary substantially across subjects for reasons unrelated to neural activity (e.g., vasodilatation), we com-
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 2C and 2F, 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...
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
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). 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).
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; Zeroldo 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 (Lou 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 frontotemporal 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 amusics 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 amusics 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 (Lou et al., 2009; Hyde et al., 2011; Albouy et al., 2013).

Critically, the ability to identify pitch-responsive regions 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 (Lou 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 neurodevelopmental impairments, so by better understanding amusia, we may gain deeper insights into the causes of neurodevelopmental disorders more broadly.

References


