Title: Does a pitch center exist in auditory cortex?

corresponding author:
Daniel Bendor
Picower Institute for Learning and Memory
Department of Brain and Cognitive Science
Massachusetts Institute of Technology
43 Vassar St., Bldg. 46, Rm. 5233
Cambridge, MA, USA.
dbendor@mit.edu

keywords: pitch, fMRI, auditory cortex, lateral Hechl’s gyrus

acknowledgements: support provided by a Charles A. King Trust postdoctoral fellowship. The author thanks Jennifer Melcher for her helpful comments and suggestions related to this manuscript.

Abstract: Pitch perception is an important component of hearing, allowing us to appreciate melodies and harmonies as well as recognize prosodic cues in speech. Multiple studies over the last decade have suggested that pitch is represented by a pitch-processing center in auditory cortex. However, recent data from Hall, Plack, and colleagues now challenge these previous claims of a human “pitch-center”.

Copyright © 2011 by the American Physiological Society.
When you pluck a guitar string, the vibrating string produces a sound composed of several spectral components—a fundamental at the frequency of the string's vibration as well as higher order harmonics at frequencies that are multiples of the fundamental. We do not hear these spectral components as separable and distinct sounds; instead they get perceptually fused together into a single sound with a pitch equal to the frequency of the fundamental. A remarkable feature of pitch perception is that it does not depend on the fundamental being physically present in the sound, as any subset of harmonics is sufficient for hearing the pitch of the “missing fundamental” (Fig 1A). Because spectrally different sounds can have the same pitch, the frequency-based neural representation of a sound in the auditory periphery does not explicitly encode pitch, and pitch must be computed at some later stage in the auditory pathway. Although there is general agreement that auditory cortex is essential for pitch perception, whether pitch processing is localized within a single functionally-specific region within auditory cortex remains a controversial issue among auditory neuroscientists.

The critical role auditory cortex plays in pitch perception was first demonstrated by Ian Whitfield more than 30 years ago [Whitfield 1980]. Whitfield trained cats to discriminate between rising and falling pitches using complex sounds where the spectrum moved in the opposite direction of the pitch. After lesioning auditory cortex bilaterally, cats were no longer able discriminate pitch changes (although frequency discrimination was still possible).

Two decades later, a region of non-primary auditory cortex (lateral Heschl’s gyrus) was identified as a pitch-processing area using fMRI imaging in human subjects [Patterson et al. 2002]. In their experiments, Patterson and colleagues used iterated ripple noise (IRN) as an acoustic stimulus, a sound created by taking a sample of broadband noise and adding this noise to itself with a delay (Fig 1B). There are two major advantages of using IRN. First, pitch salience can be parametrically changed; a weak pitch is perceived with a single delay-and-add iteration and this pitch strength increases with additional delay-and-add iterations. Next, IRN is acoustically similar to noise, except for it having a salient pitch (which noise does not). By comparing the BOLD signal change between noise and iterated ripple noise, the region of auditory cortex that responds to a sound’s pitch could be identified. Patterson and colleagues’ finding of a pitch-center in lateral Heschl’s gyrus was soon confirmed by a number of studies that collectively utilized a wide range of techniques, including fMRI, MEG, electrophysiology, and structural MRI [Penagos et al. 2004, Schneider et al. 2005, Ritter et al. 2005, Hall et al. 2006, Schönwiesner and Zatorre 2008, Puschmann et al. 2010]. Complimenting these data in humans, a putative pitch center was identified in monkeys using single-unit recordings [Bendor and Wang 2005]. The pitch center in monkeys was located in low frequency auditory cortex near the anterolateral border of primary auditory cortex, an anatomical location potentially homologous to lateral Heschl’s gyrus in humans [Bendor and Wang 2006].

Still the existence of a pitch-center in auditory cortex remains controversial. The majority of fMRI experiments identifying lateral Heschl gyrus
as a pitch center have used IRN stimuli. Given that pitch can be perceived over
a wide range of different acoustic stimuli, pitch responses should still be
observed for non-IRN acoustic stimuli. Using five spectrally different acoustic
stimuli with the same pitch, Hall and Plack (2009) observed that pitch responses
were poorly localized to one region consistently across subjects, with broad
activation mainly observed in a region posterior to lateral Heschl’s gyrus (Planum
Temporale). These data suggest that lateral Heschl’ gyrus may be activated by
some feature of IRN other than pitch. Analyzing the spectro-temporal
representation of iterated ripple noise, Hall and Plack observed modulations that
were not present in the control stimulus (Gaussian noise). Thus pitch responses
obtained by a BOLD signal subtraction (iterated ripple noise – Gaussian noise)
could correspond to acoustic differences rather the presence of a salient pitch.

Barker, Plack, and Hall (2011) examined this idea further by comparing
BOLD signal responses to IRN and an iterated ripple noise stimulus modified to
not have a pitch. Barker and colleagues created this “no-pitch” iterated noise
stimulus (IRNo) by disrupting the fine structure of iterated ripple noise, which
effectively removed the pitch while keeping the spectro-temporal modulations
that were hypothesized to be causing the pitch responses (Fig 1C). The
magnitude of these spectro-temporal modulations increased with the delay-and-
add iterations used to generate the IRN and IRNo stimuli. Across the range of
iterations used in these experiments, the spectro-temporal modulations for IRN
and IRNo stimuli were quantitatively similar, given that they produced similar
responses in a computational model of a cochlear filter bank.

Next, sixteen human subjects received fMRI scans while listening to IRN
and IRNo stimuli with 2, 4, 16, or 64 iterations, as well as two different noise
controls. Barker and colleagues observed that IRN and IRNo stimuli evoked
similar regions of activation in auditory cortex. Using a 2x4 ANOVA (p < 0.001,
corrected) no significant difference in activity between IRN and IRNo stimuli was
observed (at 2, 4, 16, and 64 iterations). Performing a similar analysis using a
region of interest (ROI) approach, this result was confirmed in central Hechl’s
gyrus, lateral Hechl’s gyrus, and Planum Temporale. All three regions of interest
did have higher activity when more delay-and-add iterations were used in
generating the IRN and IRNo stimuli. While more delay-and-add iterations
increase the pitch salience of IRN stimuli, they also increase the magnitude of
spectro-temporal modulations for both IRN and IRNo stimuli. These data
indicate that IRN activation (relative to noise) can be explained by sensitivity to
spectro-temporal modulations in the sound rather than the presence of a salient
pitch.

Next, Barker and colleagues compared the BOLD responses to IRN and
IRNo. Only 3 out of 16 subjects showed a significant difference in activity (IRN –
IRNo). Furthermore, common voxels of IRN>IRNo activation were only observed
in less than half of the subjects- in the left Heschl’s gyrus and left Planum
Temporale (6 out of 16 subjects) and in the right Planum Temporale near the
anterior border of lateral Heschl’s gyrus (7 out of 16 subjects). In contrast to this,
the activation from an IRN>noise comparison was more consistent, with 10 out of
16 subjects showing a significant effect. The regions of maximum consistency
were on the posterior edge of the right Planum Polare (14 out of 16 subjects) and in the anterolateral Planum Temporale (12 out of 16 subjects). While the IRN>noise comparison did activate lateral Heschl’s gyrus, similar to Patterson et al. (2002), the IRN>IRNo comparison did not activate this region significantly. These data suggest that IRN related activity in lateral Heschl’s gyrus is more likely a response to spectro-temporal modulations than pitch.

This study highlights the problem of using iterated ripple noise to study pitch without a proper noise control. While one interpretation of this data is that pitch responses to IRN stimuli do not occur in lateral Heschl’s gyrus, there are several issues to address before we can arrive at this conclusion. First, if a region of the brain is a pitch processing center, could it also be sensitive to non-pitch features of a sound? In monkeys, pitch-selective neurons are co-localized with frequency tuned neurons that do not respond to missing fundamental sounds [Bendor and Wang 2005]. While pitch-selective neurons are generally only found within a specific region of monkey auditory cortex, the heterogeneous composition of this “pitch-center” would make it more difficult to identify using a technique such as fMRI. Any difference between two stimuli is reflected in the BOLD signal subtraction, and therefore pitch responses could be obscured by a feature of the control stimulus unrelated to pitch that is enhanced relative to the test stimulus.

A second issue is the lack of agreement between different research groups when investigating pitch responses in lateral Heschl’s gyrus, even when using fairly similar acoustic stimuli [Penagos et al. 2004, Pushmann et al. 2010, Hall and Plack 2009]. Minor differences in acoustic stimuli or analysis methods could conceivably enhance or diminish an observed pitch response. For example, the IRN stimuli used by Barker et al. (2011) had a more limited bandwidth (1-2 kHz) than what was previously used by Patterson et al. (500-4000 Hz). Although both stimuli produced a pitch, the stimulus used by Patterson et al. had a stronger pitch salience, an attribute that has been linked to larger pitch responses in lateral Heschl’s gyrus [Penagos et al. 2004]. Secondly, the acoustic stimuli used by Barker et al. (2011) were longer in duration (580 ms) than those used by Patterson et al. (200 ms). While increasing the duration of a stimulus doesn’t change its pitch, it puts less emphasis on the onset response and more weight on the sustained response, with the latter potentially having a weaker pitch response as a consequence of neuronal adaptation.

A third point to consider is inter-subject variability. Even when group data indicates a significant pitch response in lateral Heschl's Gyrus, individual subjects can have pitch responses nearby but not directly overlapping the group data. This not unique to pitch responses, as even frequency responses to tones will give different activation maps between subjects, indicating that different fields within auditory cortex are not anchored to particular anatomical landmarks [Humphries et al. 2010]. Because imaging studies typically average group data using anatomical landmarks, inter-subject variability can obscure a significant activation pattern. This approach proved to be problematic during the initial single-unit based investigations of tonotopy in non-human auditory cortex [Evans and Whitfield 1964, Goldstein et al. 1970]. Recording a small number of neurons
from each unanaesthetized subject and grouping all the data together into an averaged frequency map using anatomical landmarks, these studies failed to demonstrate any reliable tonotopic organization in primary auditory cortex. Repeating these experiments in anesthetized subjects, which permitted the recording of a complete frequency map within a single subject, Merzenich and colleagues successfully demonstrated a reliable tonotopic organization in primary auditory cortex [Merzenich et al. 1975]. Realigning activation patterns based on the location of different physiologically defined auditory fields for group averages is more challenging, but this may be required to accurately identify pitch responses anchored to a particular auditory field. Using a similar strategy, Hall et al. (2006) recorded low and high frequency pure tone responses and pitch responses to iterated ripple noise. They observed that pitch responses overlapped low frequency responses within individual subjects. These data match observations of pitch responses in monkeys, where pitch-selective neurons were found in a low frequency region of auditory cortex.

While Hall, Plack and colleagues have not found pitch responses in lateral Heschl’s gyrus, their data does suggest the possibility of a pitch center in anterior Planum Temporale. Although adjacent to lateral Heschl’s gyrus, Planum Temporale is anatomically distinct. However, a recent mapping of tonotopicty in human auditory cortex [Humphries et al. 2010] has found two mirror symmetric tonotopic gradients, with the low frequency reversal localized to lateral Heschl’s gyrus and extending to the transverse temporal sulcus posterior to Heschl’s gyrus. Based on these data, the low frequency boundary of primary auditory cortex would be on the posterior edge of lateral Heschl’s gyrus. As such, both anterior Planum Temporale and lateral Heschl’s gyrus may be both part of low frequency primary or “primary-like” auditory cortex, making both areas a potential region of interest for finding pitch responses. It is important to point out that previous anatomical data has suggested that primary and “primary-like” auditory cortex is confined within Heschl’s gyrus [Hackett et al. 2001], and other recent frequency mapping studies (fMRI) have supported these data [DaCosta et al. 2011]. Much like the contrary results reported from pitch studies, methodological differences and intersubject variability could be responsible for this anterior-posterior shift in the physiological placement of primary auditory cortex between different fMRI studies.

In conclusion, the existence of a pitch center in lateral Heschl’s gyrus remains a contentious issue among auditory neuroscientists. Several recent published reports from Hall, Plack, and colleagues have not found evidence of a pitch center in lateral Heschl’s gyrus. These data highlight the importance of using both multiple acoustic stimuli and acoustically well-matched control stimuli to study pitch responses. Future studies would benefit from physiologically mapping auditory cortex prior to studying pitch responses, both for identifying which auditory fields contain pitch responses, and minimizing inter-subject variability within group averaged data. While a lack of pitch responses in lateral Heschl’s gyrus cannot alone negate the available data supporting a pitch center, it does suggest that more experiments are necessary to validate the existence of a pitch center.


Figure 1: Description of acoustic stimuli

A. Different subsets of harmonics will generate the same pitch. Here the addition of harmonics 1-3 and 4-6 produce complex tones with a periodicity equal to the fundamental frequency (f0, red), and are perceived as having the same pitch. The spectrum of each complex tone (green) is displayed above the time-amplitude waveform (blue).

B. Iterated ripple noise is generated by adding a noise sample to delayed version of itself (top). Repeating this add-and-delay process increases the pitch strength of the resulting sound (bottom). The delay used determines the pitch of the IRN stimulus, with a shorter delay generating a higher pitch.

C. The three acoustic stimuli used by Barker, Hall, and Plack (2011). Noise has a flat overall spectrum and weak spectro-temporal modulations. Iterated ripple noise (IRN) has a harmonic spectrum, related to its strong pitch strength and also has strong spectro-temporal modulations. To create IRNo stimuli, the fine structure of IRN stimuli is processed to effectively flatten the harmonic spectrum to look more like noise, in turn disrupting the pitch. This procedure keeps the strong spectro-temporal modulations present in the IRN stimulus intact.
Figure 1

A  Spectral invariance of pitch

harmonic #  pitch=f0  amplitude vs. frequency
6th  +  5th  +  4th
3rd  +  2nd  +  1st (f0)

B  Making iterated ripple noise

1/f0 delay

IRN

0 iterations  (no pitch)  weaker pitch salience

number of iterations  stronger pitch salience

C  Comparison of noise, IRN, and “no pitch” IRN (IRNo)

overall spectrum

spectro-temporal modulations

frequency

higher power

lower power