NeuroImage 106 (2015) 161–169

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Representation of pitch chroma by multi-peak spectral tuning in human auditory cortex

Michelle Moerel [a,*], Federico De Martino [b,c], Roberta Santoro [b,c], Essa Yacoub [a], Elia Formisano [b,c]

[a] *Department of Radiology, Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN 55455, USA*
[b] *Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, Maastricht University, Maastricht, 6200 MD, the Netherlands*
[c] *Maastricht Brain Imaging Center (MBIC), Maastricht University, Maastricht, 6229 EV, the Netherlands*

ARTICLE INFO

Article history:
Accepted 20 November 2014
Available online 3 December 2014

Keywords:
Octave percept
Human auditory cortex
Music
High field functional MRI
Pitch chroma
fMRI encoding

ABSTRACT

Musical notes played at octave intervals (i.e., having the same pitch *chroma*) are perceived as similar. This well-known perceptual phenomenon lays at the foundation of melody recognition and music perception, yet its neural underpinnings remain largely unknown to date. Using fMRI with high sensitivity and spatial resolution, we examined the contribution of multi-peak spectral tuning to the neural representation of pitch chroma in human auditory cortex in two experiments. In experiment 1, our estimation of population spectral tuning curves from the responses to natural sounds confirmed—with new data—our recent results on the existence of cortical ensemble responses finely tuned to multiple frequencies at one octave distance (Moerel et al., 2013). In experiment 2, we fitted a mathematical model consisting of a pitch chroma and height component to explain the measured fMRI responses to piano notes. This analysis revealed that the octave-tuned populations—but not other cortical populations—harbored a neural representation of musical notes according to their pitch chroma. These results indicate that responses of auditory cortical populations selectively tuned to multiple frequencies at one octave distance predict well the perceptual similarity of musical notes with the same chroma, beyond the physical (frequency) distance of notes.

Introduction

Human experience of musical notes is influenced by two relevant perceptual dimensions. The first is the *pitch height*, which relates to the physical (frequency) dimension along which notes can be ordered from low to high. The second is the *pitch chroma*, related to the similarity of musical notes at octave intervals (i.e., notes with fundamental frequency [f_0] at a 2:1 ratio; Shepard, 1982). Pitch chroma is central to melody recognition and music perception. The fundamental importance of the octave interval in perception is illustrated by the observation that it is the only interval common to nearly all musical scales ever evolved (Randel, 2003). Infants already generalize a melody across octaves (Demany and Armand, 1984), and even monkeys assign a special status to octaves, judging melodies transposed by multiples of octaves as being more similar to the original than non-integer transpositions (Wright et al., 2000). The widespread occurrence, early onset, and generalization beyond the human species of octave perception suggest that a mechanism for the explicit representation of octave frequency intervals may be present in the brain.

To date, the neural underpinnings of pitch chroma perception remain largely unknown. Throughout the auditory system, neurons are described by their characteristic frequency (CF), which is the frequency to which they respond best. Auditory neurons are spatially ordered according to their CF, resulting in tonotopic maps. This topographic organization of frequency is maintained in the various auditory relays (King and Nelken, 2009; Merzenich and Brugge, 1973; Merzenich et al., 1975) and can be examined non-invasively in the human auditory cortex using functional MRI (fMRI; Da Costa et al., 2011; Formisano et al., 2003; Humphries et al., 2010). Beyond the tonotopic organization (i.e., main peak of a voxel's spectral profile), a large part of auditory cortical neurons (Kadia and Wang, 2003) and auditory voxels' receptive fields (Moerel et al., 2013) are tuned to multiple frequency bands. In a recent study, we performed a data-driven identification of voxels with a similar pattern of multi-peaked spectral tuning and hypothesized that resulting spectral tuning patterns may be useful for auditory perception. For example, a cluster of voxels with broad spectral tuning resulting from our analysis could process overall sound energy, the harmonic tuning in another cluster could serve to parse harmonic sounds (such as speech) from a noisy background, and fine-grained tuning to multiple octaves as seen in yet another cluster ("octave cluster") may contribute to the human percept of octave similarity (Moerel et al., 2013; see Fig. 1 and below for further elaboration on this "spectral" hypothesis and a discussion of possible alternative "temporal" mechanisms).

---

* Corresponding author at: Department of Radiology, Center for Magnetic Resonance Research, University of Minnesota, 2021 Sixth Street SE, Minneapolis 55455 MN, USA. Fax: +1 612 626 2004.
E-mail address: mlmoerel@umn.edu (M. Moerel).

http://dx.doi.org/10.1016/j.neuroimage.2014.11.044

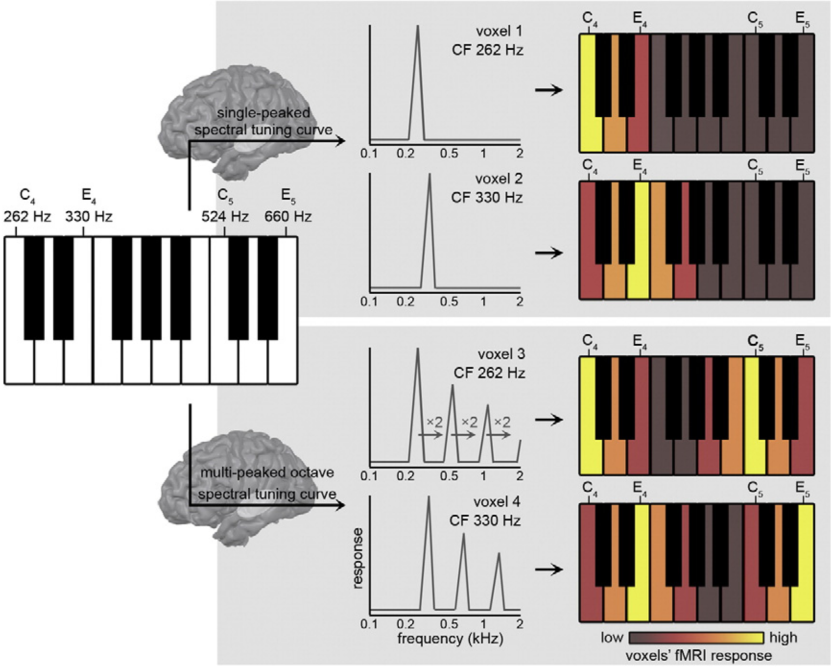

Fig. 1. Hypothesized neural mechanism for pitch chroma perception. Incoming musical sounds—such as piano notes (left)—are filtered according to the voxels' spectral tuning curves. While voxels with single-peaked spectral tuning curves only respond to musical notes that are f_0-matched with their CF (top row), octave-tuned voxels additionally respond to sounds whose f_0 matches the voxel's octave-related spectral peaks (bottom row). In this manner, octave-tuned voxels respond similar to sounds whose f_0 differs precisely one or more octaves (i.e., C$_4$ and C$_5$, or E$_4$ and E$_5$).

However, these hypotheses have yet to be tested. That is, it is not clear whether these different clusters of spectral tuning patterns, consistently observed in human auditory cortex, have any meaning from a perceptual and behavioral point of view.

Here we hypothesized that the fine-grained spectral tuning to multiple octaves in the human auditory cortex relates to the pitch chroma percept of musical notes. Specifically, we reasoned that the voxels' tuning reflects the spectral neural filtering of incoming sounds (see Fig. 1). While voxels with single-peaked spectral tuning curves only respond to musical notes that are f_0-matched with their CF (top row of Fig. 1), octave-tuned voxels are predicted to additionally display a responses to sounds whose f_0 matches the voxel's octave-related spectral peaks (bottom row of Fig. 1). For example, an octave-tuned voxel with a CF of 262 Hz (see voxel 3 in Fig. 1) will respond strongly to CF-matched piano note C$_4$, but also to notes—such as C$_5$—whose f_0 matches the voxel's octave-related spectral peaks (f_0 of C$_5$ = 524 Hz, which is 2*voxels' CF). An octave-tuned voxel with a CF of 330 Hz (see voxel 4 in Fig. 1) will respond strongly to piano notes E$_4$ and E$_5$ (with f_0 of 1*CF = 330 Hz and 2*CF = 660 Hz, respectively) and weaker to C$_4$ and C$_5$. In this manner, octave-tuned voxels respond more similar to sounds with an f_0 differing precisely one or more octaves than to sounds with an f_0 at other musical intervals, generating a representation according to the pitch chroma of these sounds.

We tested our hypothesis using data from two experiments (see Fig. 2a and b, respectively) where we used ultra-high field functional MRI (7 T) to measure auditory cortical responses to sounds. In experiment 1 (see Fig. 2a), we measured the brain's responses to a large set of natural sounds and estimated population spectral tuning curves for voxels in the sound frequency domain (i.e., population receptive fields; see top and middle row of Fig. 2a). In a data-driven manner, we identified those locations that displayed sensitivity to multiple frequency bands at one octave distance from each other (i.e., octave-tuned locations; see bottom row of Fig. 2a). In experiment 2, we measured—in the same subjects—the brain responses to piano notes using fMRI (see top row of Fig. 2b). Then we fitted the responses to piano notes (spanning three octaves) with a mathematical model (Briley et al., 2013) consisting of a pitch chroma and a pitch height component (Shepard, 1982; see middle and bottom row of Fig. 2b, and Fig. 1c of Briley et al., 2013). We expected that in the octave-tuned locations (as identified in experiment 1)—but not in other parts of auditory cortex—responses would be significantly explained by the pitch chroma component of the model, reflecting that in these locations the response to notes at a distance of one octave was more similar than the response to notes at other musical intervals.

Materials and methods

Subjects

Six subjects participated in this study (mean age [SD] = 29.5 [5.5]; three males and three females). The subjects reported to have normal hearing, had no history of hearing disorder/impairments or neurological disease, and gave informed consent before commencement of the measurements. The Institutional Review Board for human subject research at the University of Minnesota granted approval for the study.

Stimuli

The stimuli consisted of recordings of 72 natural sounds (including human speech, animal cries, musical instruments, and tool sounds) and 18 piano notes. The piano notes were sampled from an online database (University of Iowa Musical Instrument Samples; http://theremin.music.uiowa.edu/MIS.html; accessed on 11 July 2012) and consisted of six musical notes (i.e., C, D, E, G^b, A^b, and B^b) sampled from three octaves (C$_4$ = 262 Hz, C$_5$ = 523 Hz, and C$_6$ = 1047 Hz). Sounds were sampled at 16 kHz and their duration was 1000 ms. Sound onset and offset were ramped with a 10 ms linear slope, and their energy (RMS) levels were equalized. Before starting the experiment, with the earbuds in place, the subjects rated the perceived loudness of the piano notes. That is, the highest piano note (B$^b{}_6$) was played in combination with each of the other notes. The subjects adjusted the loudness of this second note until they perceived it as equally loud as B$^b{}_6$. During the experiment,