Music selectivity in the cortex is independent of extensive musical training

Dana Boebinger[1,2] (danaboebinger@g.harvard.edu)

Sam Norman-Haignere[3,4] (svnh@mit.edu)

Josh McDermott[1,2] (jhm@mit.edu)

Nancy Kanwisher[2] (ngk@mit.edu)

[1] Speech and Hearing Bioscience and Technology, Harvard University, Cambridge MA 02138 USA
[2] Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge MA 02139 USA
[3] HHMI Postdoctoral Fellow of the Life Sciences Research Foundation, Chevy Chase, MD
[4] Laboratoire des Sytèmes Perceptifs, Département d'Études Cognitives, École Normale Supérieure, PSL Research University, CNRS, Paris France

Abstract:
Speech and music perception are core functions of the human auditory system, but we know very little about the neural systems that process these complex auditory signals. While speech selectivity is an established feature of non-primary auditory cortex, clear neural selectivity for music has only recently been demonstrated (Norman-Haignere, Kanwisher, & McDermott, 2015). Here we ask whether music selectivity in the cortex requires extensive musical training, or whether it is present even in individuals with minimal musical training. To answer this question, we scanned 10 people with extensive musical training and 10 with almost none, and used the voxel decomposition methods of Norman-Haignere et al. (2015) to test whether the magnitude or anatomical extent of music-selective neural populations is influenced by musical training. Overall, we found no major differences in the patterns of fMRI responses to music stimuli between musicians and non-musicians, suggesting that music selectivity in the brain is not dependent on extensive musical training. These data raise the possibility that music selective brain responses could be a universal property of human auditory cortex.

Keywords: audition; music; fMRI; category selectivity

Introduction

As humans, some of the most complex and socially important sounds we encounter on a daily basis are speech and music. And yet we know relatively little about the neural systems that process these sound categories. Based on a large body of prior research, we know much about tuning within human primary auditory cortex for simple acoustic features like frequency (Humphries, Liebenthal, & Binder, 2010) and spectrotemporal modulation (Santoro et al., 2014; Schönwiesner & Zatorre, 2009). More recent work has focused on understanding the tuning properties of non-primary regions of human auditory cortex, finding evidence for neural populations that are selective for a particular category of sound, such as speech or music (Angulo-Perkins et al., 2014; Norman-Haignere et al., 2015; Staeren, Renvall, De Martino, Goebel, & Formisano. 2009).

However, it is unknown whether these stimulus selectivities are equally present in all individuals, or whether they are dependent on or modulated by training. Musical training has been shown to be associated with widespread structural (Bermudez & Zatorre, 2005; Bermudez, Lerch, Evans, & Zatorre, 2009; Gaser & Schlaug, 2003; Kleber et al., 2016) and functional (Angulo-Perkins et al., 2014; Margulis et al., 2009; Pantev et al., 2001) changes to cortex.

In this study, we used the voxel decomposition method introduced by Norman-Haignere et al. (2015) to test whether the cortical music selectivity found in that study arises only (or to a greater degree) in those with extensive musical training, or whether it is a widespread property of normal human brains. We also tested whether stimulus categories like drums and non-Western music produce a different response from other Western musical sounds in these music-selective neural populations.

Methods

Twenty young adults (14 female, mean = 24.7 years) participated in the experiment: 10 musicians (8 female, mean = 23.5 years) and 10 non-musicians (6 female, mean = 25.8 years). Inclusion criteria for musicians included beginning formal training before the age of seven, and continuing training until the current day. To be classified as a non-musician, subjects were required to have less than two years of total music training, which could not have occurred either before the age of seven or within the last five years. There were no significant group differences in age, education, or socioeconomic status. All subjects were native English speakers and had normal hearing (audiometric thresholds <25 dB HL for octave frequencies 250Hz to 8kHz).

Each subject underwent a 2-hour behavioral testing session as well as three 2-hour fMRI scanning sessions. During the behavioral session, subjects completed an audiogram, filled out questionnaires about their musical experience, and completed a series of basic psychoacoustic tasks.

During the scanning sessions, sounds were presented in a "mini-block design," in which each 2-second natural sound was repeated three times, with a single fMRI volume collected between each repetition (i.e. "sparse scanning"). To encourage subjects to pay attention to the sounds, either the second or third repetition in each "mini-block" was 12dB quieter, and subjects were instructed to press a button when they heard this quieter sound. Overall, subjects performed well on this task (>85% average performance across runs).

Stimuli consisted of 2-second clips of 192 different natural sounds. These sounds included the 165 stimuli used in Norman-Haignere et al. (2015), which included some of the most frequently heard and recognizable sounds in everyday life, as well as additional music and drumming clips from a variety of musical cultures.

MRI Data were collected at the Athinoula A. Martinos Imaging Center of the McGovern Institute for Brain Research at MIT, on a 3T Siemens Prisma, TR = 3.4s, TA = 1s, whole brain, voxel size: 2 x 2 x 2.8 mm (10% gap), 46 Slices, FoV: 192 mm (96 x 96 matrix).

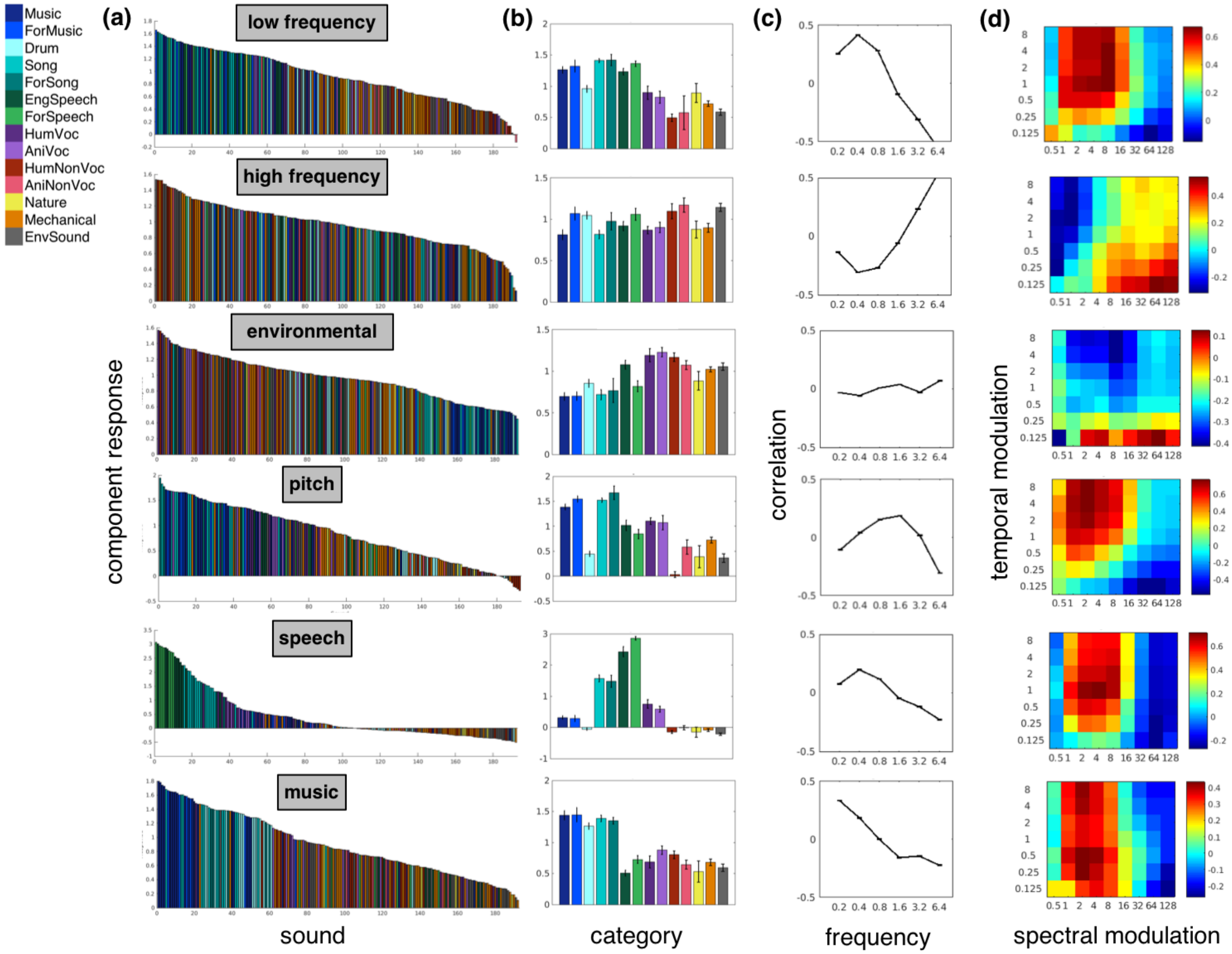

Figure 1. (a) Response profiles inferred for each component for all 192 sounds tested. Sounds are ordered by magnitude and colored by semantic category. **(b)** Component responses averaged across sounds from the same category. **(c)** Correlation of component response profiles with energy in different frequency bands. **(d)** Correlation of component response profiles with spectrotemporal modulation energy in the cochleograms for each sound.

Results

Behavioral results showed that musicians were significantly better at discriminating melodies (t(18) = 4.22, p = 0.0005), detecting violations within melodies (t(18) = 5.27, p = 0.0001), synchronizing with an isochronous beat (t(18) = -3.48, p = 0.0027), and had marginally better frequency discrimination thresholds (t(18) = 1.82, p = 0.0860).

As in Norman-Haignere et al. (2015), we modeled the response of each voxel as a weighted sum of canonical response patterns to the sound set ("components"). Our results replicate the results from Norman-Haignere et al (2015) as follows. Six components were found to explain over 80% of the replicable variance in the responses across auditory cortex. Four of these components reflected selectivities for acoustic features of the sounds (e.g. frequency, spectrotemporal modulation), and two showed selectivity for the high-level categories of music and speech (Figure 1).

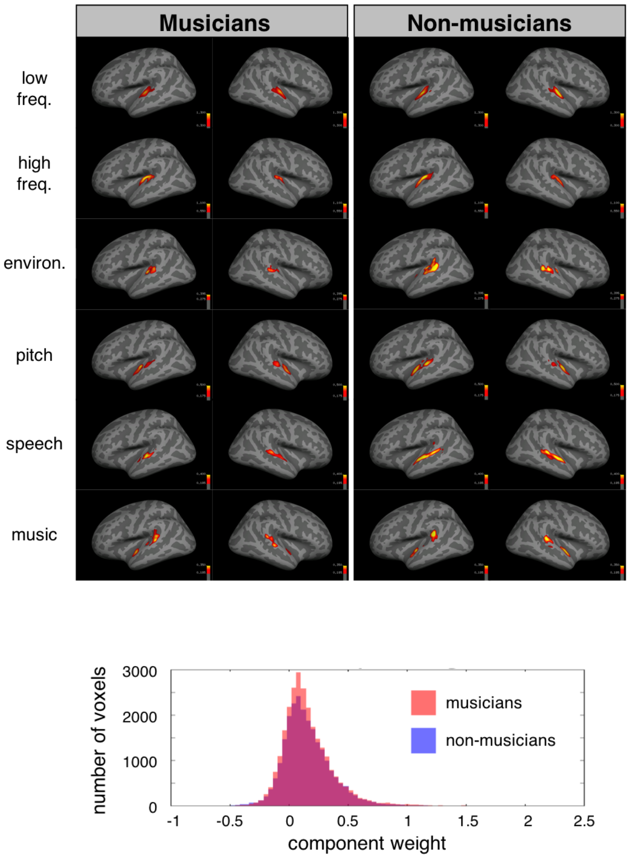

Figure 2. (top) Component voxel weights, averaged across subjects separately for musicians (left) and non-musicians (right). **(bottom)** Histograms showing music component weights for each group.

When these speech- and music-selective components were projected back onto the brain, they concentrated in non-overlapping regions of non-primary auditory cortex, and their anatomical distribution was similar across subjects.

We then assessed the extent to which musical training affects stimulus selectivity in auditory cortex by comparing the component weights between expert musicians and non-musicians in both their magnitude and anatomical distribution. We found no apparent difference between musicians and non-musicians in any of these measures (Figure 2). Histograms of the music component voxel weights were very similar, but did differ significantly between groups in their medians (musicians = 0.1213, non-musicians = 0.1154, Z = 2.1559, p =0.0031), and skew (musicians = 0.9244, non-musicians = 1.2168, p < 0.0001, 10,000 permutations). These group differences were comparable to those for all other components, and are likely due to the small number of subjects included in this study.

When the same voxel decomposition analysis was repeated for musicians and non-musicians separately, the resulting response profiles for the two groups were correlated with each other with a mean of r = 0.8395. This suggests that similar component structures underlie auditory cortical responses in both groups (Figure 3), though further research will investigate whether the differences between these component structures is interesting or meaningful.

The expanded stimulus set used in this study allowed us to gain further insight into the music-selective component discovered in Norman-Haignere et al. (2015). This component has high weights for both non-Western musical and drum stimuli, suggesting a fairly general function for the underlying neural computations.

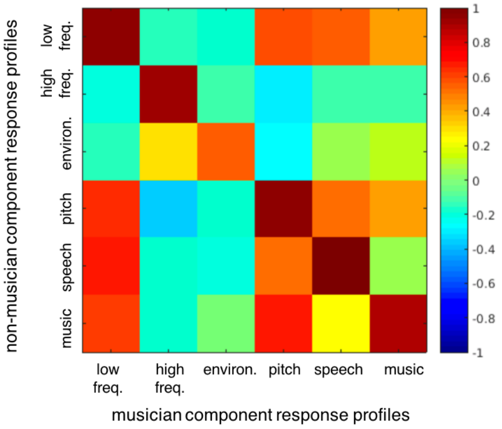

Figure 3. Correlation matrix (Pearson's R) between response profiles inferred from musicians and non-musicians separately.

Discussion

We used hypothesis-free voxel decomposition methods (Norman-Haignere et al., 2015) to uncover music-selective neural populations in non-primary auditory cortex, and examined whether these neural populations were affected by musical experience. Results show no significant difference between the magnitude, selectivity, or anatomical extent of these neural populations in expert musicians compared to individuals with no musical training This suggests that music selectivity is a fundamental characteristic of the functional organization of human non-primary auditory cortex, not dependent on extensive musical training.

Of course, even our non-musician participants had extensive exposure to music, and given the universality of music in human cultures, it would be difficult to find otherwise normal humans who lacked exposure to music. Thus, our data do not show that music selectivity in the brain is independent of experience. Rather, our study shows that normal exposure to music is sufficient for music selectivity in the brain, without extensive musician training.

Acknowledgments

This work was supported by NSF grant BCS-1634050 to J.M. and NIH grant DP1HD091947 to N.K.

References

Angulo-Perkins, A., Aub E, W., Peretz, I., Barrios, F. A., Armony, J. L., & Concha, L. (2014). Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians. *Cortex*, *59*, 126–137.

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences, 1060, 395-399.*

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex, 19*(7), 1583-1596.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and nonmusicians. *Journal of Neuroscience, 23*(27), 9240-9245.

Humphries, C., Liebenthal, E., and Binder, J.R. (2010). Tonotopic organization of human auditory cortex. *NeuroImage, 50*(3)*,* 1202–1211.

Kleber, B., Veit, R., Valérie, C., Gaser, C., Birbaumer, N., & Lotze, M. (2016). Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage*, *133*, 477–483.

Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B., & Wong, P. C. M. (2009). Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Human Brain Mapping*, *30*(1), 267–275.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, *88*(6), 1281–1296.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport, 12*(22), 169-174.

Santoro, R., Moerel, M., De Martino, F., Goebel, R., Ugurbil, K., Yacoub, E., & Formisano, E. (2014). Encoding of Natural Sounds at Multiple Spectral and Temporal Resolutions in the Human Auditory Cortex. *PLoS Computational Biology, 10*(1).

Schönwiesner, M., & Zatorre, R. J. (2009). Spectrotemporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proceedings of the National Academy of Sciences, 106*(34), 14611-14616.

Staeren, N. L., Renvall, H., De Martino, F., Goebel, R., & Formisano, E. (2009). Sound categories are represented as distributed patterns in the human auditory cortex. *Current Biology, 19*, 498–502.