

Effects of Asymmetric Cultural Experiences on the Auditory Pathway

Evidence from Music

Patrick C. M. Wong,^a Tyler K. Perrachione,^b
and Elizabeth Hellmuth Margulis^c

^a*The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders and Department of Otolaryngology – Head and Neck Surgery, Northwestern University, Evanston/Chicago, Illinois, USA*

^b*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*

^c*Department of Music, University of Arkansas, Fayetteville, Arkansas, USA*

Cultural experiences come in many different forms, such as immersion in a particular linguistic community, exposure to faces of people with different racial backgrounds, or repeated encounters with music of a particular tradition. In most circumstances, these cultural experiences are asymmetric, meaning one type of experience occurs more frequently than other types (e.g., a person raised in India will likely encounter the Indian *todi* scale more so than a Westerner). In this paper, we will discuss recent findings from our laboratories that reveal the impact of short- and long-term asymmetric musical experiences on how the nervous system responds to complex sounds. We will discuss experiments examining how musical experience may facilitate the learning of a tone language, how musicians develop neural circuitries that are sensitive to musical melodies played on their instrument of expertise, and how even everyday listeners who have little formal training are particularly sensitive to music of their own culture(s). An understanding of these cultural asymmetries is useful in formulating a more comprehensive model of auditory perceptual expertise that considers how experiences shape auditory skill levels. Such a model has the potential to aid in the development of rehabilitation programs for the efficacious treatment of neurologic impairments.

Key words: bimusicality; music cognition; neural correlates; language; fMRI

Introduction

Music is a significant part of human culture, occurring not only in concert halls and opera houses, but also throughout the daily environment on street corners, in shops, and on radios, iPods, televisions, and computers. The impact of formal musical training on brain anatomy^{1,2} and physiology³ has sparked much interest in scientific communities and the general public. Formal musical training is typically (though

not exclusively) conceived as extensive specialized training in a particular musical instrument, played in the style of a particular cultural tradition, such as training as a pianist in a conservatory. The term “musician” is often used to refer to an individual who has engaged in such lengthy, explicit, and formal training. However, an important aspect of music is that it serves a variety of functions in the daily lives of people from all cultures, and is not exclusively a privilege of those who happen to receive formal training. Similarly, language is not a phenomenon that is significant only to linguists or those who study in special language schools. In our discussion of musical training and its

Address for correspondence: Patrick C. M. Wong, 2240 Campus Drive, Evanston, IL 60208. Voice: +1-847-491-2416; fax: +1-847-491-2429. pwong@northwestern.edu

impact below, we expand our definition of “musical training” to include not only formal and explicit training, but also informal and implicit training. The latter includes musical exposure from the ambient environment, which very often involves listening to music of a particular culture without explicit music making; listeners of this type are still affected by such exposure in a systematic and tractable fashion. Our broad definition of musicianship here is in line with what has been elegantly argued by Bigand and Poulin-Charronnat,⁴ who contend that there is sometimes a lack of fundamental distinction between individuals with and without formal musical training.

Our paper is divided into three sets of studies that were performed in our laboratories. First, we will focus on the impact of formal musical training on speech learning to show how extensive and explicit musical instruction can result in more efficacious speech learning, which is associated with certain neuroanatomic and neurophysiologic signatures.^{5–7} We will then focus on how extensive musical training results in a cortical network of expertise by discussing a study in which violinists and flutists listened to their instrument of expertise during fMRI scanning.⁸ Finally, we will discuss a series of experiments where we examined the impact of implicit and informal musical exposure on everyday music listeners, including those who have exposure to music of one or two cultures.⁹

Formal Musical Training and Speech Learning

One compelling way to investigate the basis for differences in behavioral expertise is to quantify anatomic and physiologic differences in the neural systems supporting those behaviors. Neural systems exhibit remarkable plasticity, and owe a large part of their development to the environmental experiences of the organism. For the human auditory system, far and away the greatest environmental influences

involve the experience of culturally meaningful language (speech) and music. The effects of asymmetric experience with differences in spoken language have been extensively documented. Individuals exhibit privileged patterns of familiarity and sensitivity to native speech sounds compared to foreign sounds, which can make the acquisition of a foreign language difficult.¹⁰ Differences in auditory abilities between native and non-native speech sounds are reflected in different patterns of neural activity with respect to these sounds.^{11,12} Empirical work has demonstrated that, despite these difficulties, native-like performance in the production and perception of non-native speech sounds can be achieved in adulthood.^{13,14}

Although the cognitive science and neuroscience literatures have long debated to what extent linguistic abilities, including speech perception, constitute a unique domain, there is substantial evidence that the neural systems responsible for all types of auditory perception and behavior are highly integrated. Behaviorally, the language you speak can have an impact on your ability to identify individuals from their voice.¹⁵ Neurally, the various regions of auditory cortex appear to participate in a variety of different auditory tasks, including both speech and music.¹⁶ Current results from neuroscience are beginning to reveal the ways in which musical expertise may facilitate spoken language learning in adulthood.

Much of the work looking at the relationship between music and speech processing has focused on the perception and learning of lexical tones. In tone languages, such as Mandarin Chinese, words are differentiated not only by their vowels and consonants, but also by their pitch patterns. In Mandarin, a word spoken with a high-level tone (shī, “teacher”) versus a falling tone (shì, “city”) can have meanings as different as words spoken with different vowels (e.g., “bike” versus “book”). Because one component of musical expertise is the ability to differentiate among many different pitches (notes), it is a reasonable hypothesis that musical expertise may also confer some advantage in

learning languages that use lexical tones. This hypothesis is furthered by the results of a study showing that subjects with formal musical training are better able to identify and discriminate the pitch contours in Mandarin lexical tones than subjects without such training.¹⁷

To better understand the relationship between musicianship, perceptual expertise, and speech learning, we trained a number of young native English-speaking adults on an artificial vocabulary in which words were distinguished by lexical tones.⁶ Of the nine subjects who successfully mastered this vocabulary, eight had more than 6 years of formal musical training. Of the seven subjects who did not fully master the vocabulary, only one reported that level of musical training. Additionally, the level of vocabulary mastery could be significantly predicted by individuals' performance on a pitch-pattern identification task—a task at which musicians significantly outperformed their non-musician peers. Even before training began, differences in the patterns of auditory neural activity, as measured by fMRI during a pitch-pattern discrimination task, distinguished subjects who would successfully master the vocabulary from those who would not.⁵ Successful learners exhibited greater activity bilaterally in regions of the posterior superior temporal lobe associated with sound-pattern classification. Less-successful learners, meanwhile, exhibited more activity in regions of the frontal lobe associated with attention and decision making, such as the anterior cingulate. After training on the lexical-tone-based vocabulary, successful learners showed greater activation in the left posterior superior temporal gyrus, which is consistent with patterns of neural activity in native tone-language speakers during a similar task.¹⁸ Less-successful learners, meanwhile, again demonstrated greater activation in a diffuse network of frontal regions associated with attention and decision making.

The differences between successful and less-successful lexical-tone learners are not limited to their behavior and neurophysiology. Gross differences in neural anatomy were also evident

between these two groups. The first region of isocortex to receive auditory connections from the thalamus is called primary auditory cortex and resides almost entirely on Heschl's gyrus (also called the transverse temporal gyrus), which is found bilaterally in the posterior superior temporal lobe. The volumes of grey matter (neuron cell bodies) and white matter (neuronal axons) in these gyri were measured in each subject who completed the vocabulary-learning paradigm described above.⁷ Compared to the less-successful learners, the successful learners showed significantly larger volumes of both grey and white matter in the left Heschl's gyrus. There was a significant positive correlation between the volume of grey matter in left Heschl's gyrus and level of vocabulary mastery. It is probable that the larger size of this gyrus represents a larger network of neurons that contribute computationally to tasks supported by this structure, which is reflected in the enhanced performance of the successful, more musically trained, learners. It is worth pointing out that, in addition to pitch, the larger anatomic volume of this region has also been associated with successful learning of non-native consonants.¹⁹

Heschl's gyrus and primary auditory cortex are the first cortical regions involved in auditory processing, yet there are a number of subcortical regions in the thalamus and brain stem that receive and process the signal prior to its arrival in cortex. Although these regions are considered "lower-level," especially with regard to complex human behaviors, such as speech and music, recent evidence suggests that differences in expertise for these behaviors are, in fact, associated with processing differences at the level of the brain stem. In one such study, the frequency-following response (FFR) was measured in musicians and nonmusicians who were passively hearing lexical tones.²⁰ The FFR is an evoked electrical potential which evidence suggests arises from the synchronous activity of neurons in the rostral brain stem (inferior colliculus and lateral lemniscus), and which tracks the fundamental frequency (here, pitch)

of an incoming auditory signal. Compared to non-musicians, the musicians exhibited a significantly higher fidelity in their FFR, indicating more accurate tracking of the pitch patterns of the lexical tones. This difference was more pronounced as the pitch patterns became more complicated (i.e., rising and dipping tones). Although a genetic basis for this result cannot be ruled out, results do suggest an important role of experience: A positive correlation between years of musical training and FFR fidelity indicated that the longer one trained musically, the more robust the brain stem encoding of pitch was. Similarly, a negative correlation between onset of musical training and FFR fidelity indicated that the earlier in life one started musical training, the stronger the neural representation of pitch was. Subsequent work has indicated that explicit training on the use of lexical tones can, in fact, improve the fidelity of brain stem encoding of pitch in adults.²¹

Taken together, these studies provide compelling data to believe that the processing of music and speech may rely, at least in part, on a shared neural architecture. Musicians exhibit enhanced encoding of linguistic pitch at the level of the brain stem.²⁰ Successful learners of an artificial tone-language vocabulary tend to be predominately those with musical training⁶; and, not only do these successful learners display patterns of cortical activity more like native speakers of a tone language,⁵ but they also display a larger volume in the macroanatomic cortical structure associated with linguistic pitch processing.⁷ Although the exact mechanisms by which experience with music and speech serves to underwrite the same neural networks remain an important area of research, the evidence suggests there does exist a meaningful, facilitatory relationship between these two complex auditory behaviors.

Formal Musical Training and a Neural Networks of Expertise

Formal musical training involves a special set of long-term experiences with the sound of a

performer's instrument of expertise, including experience creating, evaluating, and listening to these sounds. People with formal musical training have been shown to recruit distinct areas for music processing³; however, studies on this subject generally contrast participants with and without formal musical training, leaving open a potential explanation rooted in different genetic predispositions (for example, musical training may be sought out by people with a genetic predisposition, and avoided by people without it). The case for the impact of training on the development of these distinct neural responses is strengthened by a recent study, which contrasted two subject groups (one group of people with extensive formal training on the violin, and another of people with extensive training on the flute) who share the experience of significant training in the classical music tradition.⁸ Insofar as it is unlikely that genetics may predispose someone to play the flute versus the violin (although it is plausible that they may predispose someone to study music in the first place), it is unlikely that the study's results stem from genetic factors, and more likely that they represent an effect of training.

In the study, violinists and flutists listened to short style-matched excerpts (from Bach Partitas) for flute and violin while being scanned with fMRI. A comparison of responses to the instrument of expertise (flute for flutists, violin for violinists) with responses to the other instrument (violin for flutists, flute for violinists) revealed an extensive cerebral network of expertise, encompassing regions devoted to the processing of musical syntax (BA 44), timbre (auditory association cortex), self-relevance (frontal regions), and motor planning (precentral gyrus). Figure 1, for example, shows the activation in the precentral gyrus and BA 44 elicited by stimuli on each instrument in each group. Activity in these areas was robust when each group heard stimuli played on their own instrument, and comparatively weak when they heard stimuli played on the other one. That people could respond so differently to music similar in every way except for its instrument reveals that

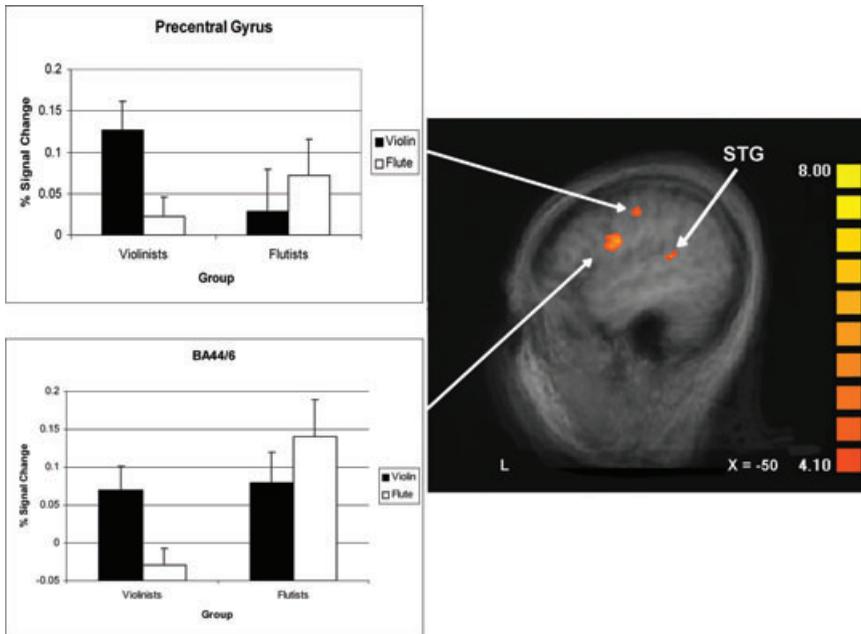


Figure 1. Brain activation revealed by contrasting instrument of expertise versus instrument of nonexpertise (based on a random effect analysis) showing activation in left BA 44/6, precentral gyrus, and STG. Bar graphs show activation for each instrument for each subject group. Error bars indicate standard error of the mean. Activation is projected onto a T1-weighted volume averaged across all subjects; color bar indicates strength of activation in *t*-value.⁸ (In color in *Annals* online.)

musical responses are shaped not by structure alone, but also by personal experience and listening histories. These results are consistent with studies showing that musical training leads to preferentially enhanced cortical representations for authentic versus synthetic timbres only on an instrument of expertise.²² Long-term exposure and active use have the power to reshape responses to auditory stimuli.

Everyday Music Listening and Bimusicality

Although one special type of musical experience involves explicit formal training and active involvement playing an instrument, a far more common type involves implicit training (through exposure to music in everyday life) and more passive involvement (listening rather than playing). This kind of training is clearly effective; the everyday mu-

sic listener can be passionately moved by his or her favorite songs, despite no formal training.

Culture constrains the exposure listeners receive: Western listeners, for example, are more likely to encounter Western than Indian music. This situation parallels that of language, where speakers are enculturated into English or Chinese, for example, by virtue of exposure to a particular environment. Bilingualism is a common occurrence, whereby individuals learn to speak two or more languages with varying levels of proficiency in different social contexts. But language experience includes not only extensive listening, but also extensive production experience with speaking the language. We have asked whether “bimusicality” can arise in response to extensive exposure to music from two cultures, even without an active production component (that is, without experience playing an instrument or otherwise creating music), such as exists in bilingualism.⁹

In the study, three groups of everyday adult listeners without significant formal training participated in two experiments assessing cognitive and affective responses to excerpts of Western and Indian music. A recognition task was used to assess cognitive responses. Participants heard 30-s excerpts of Western music (symphonies by J. Stammitz and G.B. Sammartini) and Indian music (compositions by N. Banerjee and U.R. Skhan). Later, they were presented with 4- to 6-s clips from the same works, some of which had been heard and some of which had not, and were asked whether they had heard them in the exposure phase. A tension judgment task was used to assess affective responses. Participants were presented with 10- to 18-s Western and Indian melodies matched for tempo, tonic pitch, and timbre (half were played on the piano, half on the sitar), but differentiated by scale and meter. Western excerpts used major and minor scales, and Indian excerpts used bhairav and todi. Similarly, Western excerpts used 2/4, 3/4, and 4/4 meters, but Indian excerpts used tinal, ektal, and rupaktal.

Both the cognitive and affective tasks provided evidence of bimusicality, with monomusical Western and Indian listeners showing differentiated responses for music from the two cultures, and bimusical listeners who had been enculturated into both Western and Indian musical traditions failing to exhibit a differentiated response between the two. For example, while Western listeners perceived more tension in Indian excerpts, and Indian listeners perceived more tension in Western excerpts, the bimusical group made no such distinction. In terms of recognition memory, Western listeners were more accurate for Western excerpts, and Indian listeners more accurate for Indian excerpts, but bimusical listeners were not significantly more accurate with recognition memory for either stimulus type. Ongoing neuroimaging work contrasting monomusical and bimusical listeners seeks to understand the circuitry underlying bimusicality, in particular, to understand whether there is a difference between

early bimusicals, who receive exposure during a critical window in early childhood, and late bimusicals, who do not receive exposure until later in adulthood. It is notable that the bimusicality revealed in the behavioral studies arose in response to relatively passive exposure, without the active-use component characteristic of bilingualism. People can acquire sensitivities to complex auditory stimuli associated with multiple cultures simply through exposure and enculturation, without the kind of intensive participation and activity characteristic of actually speaking a language. Listening to music is enough.

Conclusion

We have presented studies that show the impact of both formal and informal musical training/exposure on our nervous system and on our cognitive and affective behaviors. The fact that both types of training/exposure show overlapping and unifying characteristics, and the fact that they both extensively permeate our nervous system, arguably provides fundamental justifications for the existence of auditory processing disorders²³ as a class of disorders as well as the promise of behavioral interventions. Although some commonalities could be isolated, we also see variability and individual differences. To a certain extent, neural signatures that were associated with ultimate learning success could be identified even before training. In discussing music as a cultural phenomenon, and with the consideration of individual differences, it would be insufficient not to acknowledge the fact that cultural differences exist in domains of human behaviors,²⁴ although little emphasis has been placed on auditory behaviors. Future research will need to seriously consider possible cultural differences in the auditory domain. It is our hope that by examining formal training and informal (cultural) influences in auditory perception and learning, we will gain a more complete understanding of the auditory system and will be able to develop

clinical practices that can serve people of all cultures.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Gaser, C. & G. Schlaug. 2003. Brain structures differ between musicians and non-musicians. *J. Neurosci.* **23**: 9240–9245.
- Schneider, P. et al. 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* **8**: 1241–1247.
- Gaab, N. & G. Schlaug. 2003. The effect of musicianship on pitch memory in performance matched groups. *NeuroReport* **14**: 2291–2295.
- Bigand, E. & B. Poulin-Charronnat. 2006. Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition* **100**: 100–130.
- Wong, P.C.M., T.K. Perrachione & T.B. Parrish. 2007. Neural characteristics of successful and less successful speech and word learning in adults. *Hum. Brain Mapp.* **28**: 995–1006.
- Wong, P.C.M. & T.K. Perrachione. 2007. Learning pitch patterns in lexical identification by native English-speaking adults. *Appl. Psycholinguist.* **28**: 565–585.
- Wong, P.C.M. et al. 2008. Volume of left Heschl's gyrus and linguistic pitch learning. *Cereb. Cortex* **18**: 828–836.
- Margulis, E.H. et al. 2009. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* **30**: 267–275.
- Roy, A.K., E.H. Margulis & P.C.M. Wong. 2008. Bimusicality: a dual enculturation effect on non-musicians' musical tension and memory. Presented at the Tenth International Conference on Music Perception and Cognition. Sapporo, Japan, August 25–29.
- Iverson, P. et al. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* **87**: B47–B57.
- Näätänen, R. et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **385**: 432–434.
- Wong, P.C.M. et al. 2004. The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. *J. Neurosci.* **24**: 9153–9160.
- Jamieson, D.G. & D.E. Morosan. 1986. Training non-native speech contrasts in adults: acquisition of the English /ð/-/θ/ by francophones. *Percept. Psychophys.* **40**: 205–215.
- Bradlow, A.R. et al. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *J. Acoust. Soc. Am.* **101**: 2299–2310.
- Perrachione, T.K. & P.C.M. Wong. 2007. Learning to recognize speakers of a non-native language: implications for the functional organization of human auditory cortex. *Neuropsychologia* **45**: 1899–1910.
- Price, C., G. Thierry & T. Griffiths. 2005. Speech-specific auditory processing: where is it? *Trends Cogn. Sci.* **9**: 271–276.
- Alexander, J., P.C.M. Wong & A. Bradlow. 2005. Lexical tones perception in musicians and nonmusicians. Presented at Interspeech 2005 – Eurospeech – 9th European Conference on Speech Communication and Technology. Lisbon, Portugal, September 4–8.
- Xu, Y. et al. 2006. Activation of the left planum temporale in pitch processing is shaped by language experience. *Hum. Brain Mapp.* **27**: 173–183.
- Golestani, N. & C. Pallier. 2007. Anatomical correlates of foreign speech sound production. *Cereb. Cortex* **17**: 929–934.
- Wong, P.C.M. et al. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* **10**: 420–422.
- Song, J. et al. 2008. Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cogn. Neurosci.* **20**: 1892–1902.
- Pantev, C. et al. 2001. Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport* **12**: 169–174.
- Bellis, T.J. 2003. *Assessment and Management of Central Auditory Processing Disorders in the Educational Setting*, rev. ed. Singular/Delmar Learning, Clifton Park, NY.
- Nisbett, R.E. & Y. Miyamoto. 2005. The influence of culture: holistic versus analytic perception. *Trends Cogn. Sci.* **9**: 467–473.