University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

2021

Individual auditory categorization abilities are shaped by intrinsic and experience-driven neural factors

Kelsey Mankel

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
Mankel, Kelsey, "Individual auditory categorization abilities are shaped by intrinsic and experience-driven neural factors" (2021). *Electronic Theses and Dissertations*. 2661.
https://digitalcommons.memphis.edu/etd/2661

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

INDIVIDUAL AUDITORY CATEGORIZATION ABILITIES ARE SHAPED BY INTRINSIC AND EXPERIENCE-DRIVEN NEURAL FACTORS

by

Kelsey Mankel

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Communication Sciences and Disorders

The University of Memphis

August 2021

Dedication

To my family, thank you for your endless love and encouragement in the pursuit of my dreams. To my friends, professors, and mentors, I am grateful for your support and thoughtful advice that guided my academic journey. And in memory of my friend and teacher Dan Ruggles, who fostered my passion for music and taught me to appreciate the beautiful sounds and people around me. Thanks to you, I will always strive to be “all that and a bag of chips.”

Acknowledgements

I would like to extend my deepest appreciation to Dr. Gavin Bidelman for his wonderful mentorship and support. From the beginning, he saw my greatest potential and helped me charter a path to achieve it, gently pushing when I needed it to reach for higher goals than I thought possible. I know my future will be bright and successful if I have gained even a fraction of his drive, passion, and wisdom during my PhD experience. I will be forever grateful for his guidance over the past five years.

I also sincerely thank my committee members Dr. Eugene Buder, Dr. Deborah Moncrieff, and Dr. Philip Pavlik, Jr. Their questions and suggestions motivated me to think more critically about my dissertation topic, approach the problem from a new angle, or broaden my scope to address important issues in human auditory health.

Finally, thank you to my friends and family for their support throughout this PhD experience. I am grateful for my lab mates, fellow CSD PhDs, and IISSO colleagues for our stimulating discussions and friendships. The completion of this dissertation would not have been possible without them.

This work was supported by the National Institute on Deafness and Other Communication Disorders from the National Institutes of Health under award numbers F31DC019041 (KM) and R01DC016267 (GMB). This work was also supported by the Institute for Intelligent Systems Student Organization Dissertation Award at the University of Memphis.

Preface

Chapter 2 was published as a manuscript in *NeuroReport*. Chapter 3 has been submitted for publication and is currently under review at the *Journal of Cognitive Neuroscience*. It has also been posted as a preprint to *bioRxiv*.

Mankel, K., Barber, J., & Bidelman, G. M. (2020). Auditory categorical processing for speech is modulated by inherent musical listening skills. *NeuroReport*, *31*(2), 162-166. https://doi.org/10.1097/WNR.0000000000001369

Mankel, K., Shrestha, U., Tipirneni-Sajja, A., & Bidelman, G. M. (2021). Functional plasticity coupled with structural predispositions in auditory cortex shape successful music category learning. *bioRxiv [preprint]*. https://doi.org/10.1101/2021.05.12.443818

Abstract

To make sense of the auditory world, listeners must organize diverse, continuously varying sounds into meaningful perceptual categories. The auditory categorization process is believed to be a foundational skill for language development and speech perception. Despite decades of behavioral research, neuroscientific evidence is only beginning to uncover where, when, and how auditory categories arise in the brain. Although it has been proposed that categorical perception is shaped by both innate (nature) and experience-driven (nurture) factors, it is unclear how these features manifest neurally at the individual level. In the first study of this dissertation, we recorded multi-channel electroencephalography (EEG) in nonmusicians who varied in their intrinsic musical listening skills (i.e., musicality) and evaluated their performance on a speech categorization task. The results demonstrated that listeners with naturally superior musicality exhibit sensory processing advantages within the right auditory cortex which enables more efficient and robust categorization. The second study then assessed whether neural encoding differences influenced learning of unfamiliar auditory categories. Using a rapid learning paradigm, musically naïve listeners were trained to identify musical interval categories (i.e., minor and major thirds). More successful category learners showed more efficient post-training sensory encoding of musical intervals by ~150-200 ms (i.e., P2). Structural magnetic resonance imaging (MRI) also revealed differences in gray matter and cortical thickness within auditory cortex associated with categorization performance, indicating that the acquisition of auditory categories is driven by a layering of preexisting and short-term plastic changes in brain function. Functional asymmetries were also observed across both studies, suggesting that networks in the right auditory cortex may contribute to enhanced categorization and/or may be more sensitive to individual perceptual differences in general. Taken together, these results highlight the complex interplay of both nature and nurture in auditory categorization. Intrinsic, individual differences in neural function cultivate sensory processing advantages within auditory cortex that are shaped by experience (e.g., learning) and promote enhanced categorization and perception of sounds.

Table of Contents

List of Figures

Chapter 1

General Introduction

Sensory information in the environment is varies continuously. To aid perception, the brain organizes this incredibly diverse and complex information into discrete groups. Stimuli that share similar perceptual features are classified into distinct categories where discrimination of objects is easier between different categories (e.g., /b/ vs. /d/ or blue vs. green) than among stimuli within the same category (e.g., two different instances of a /b/ sound or shades of blue). This phenomenon is known as categorical perception (Harnad, 1987).

Within the auditory realm, categorizing speech sounds (i.e., phonemes) is a foundational skill for speech perception and language development (Pisoni & Luce, 1987). For example, categorical perception allows a listener to transcend superficial differences between sounds, thus increasing perceptual resiliency to contextual influences such as different talkers, speaking rates, or fundamental frequencies (Goldstone & Hendrickson, 2010). Acquisition of speech categories begins early in life as babies are exposed to sounds of their native language (Kuhl, 1991; Kuhl, 2004; Kuhl et al., 2008). Categorization deficits have been observed in populations with reading and language difficulties, including disorders such as dyslexia and autism (Coady et al., 2005; McMurray et al., 2010; Mody et al., 1997; Noordenbos & Serniclaes, 2015; Wang et al., 2017; Werker & Tees, 1987). These results highlight a critical link between categorization abilities and speech-language skills.

Theoretical basis of categorical perception

Several theories have emerged for how perceptual objects are organized and classified in the mind, particularly how this action leads to perception of speech (Harnad, 1987). According to the motor theory of speech perception, for example, categorical perception of speech phonemes is mediated by prediction of the motor patterns that produce a speaker’s articulatory gestures

(Liberman et al., 1967; Liberman et al., 1957; Liberman & Mattingly, 1989). In contrast, auditory-based accounts propose that categorical representations arise from sensitivities within the auditory system tuned to specific acoustic properties of the signal (Klatt, 1979; Pastore, 1976; Pastore et al., 1977; Stevens, 1981). Prototype theories propose that a “perceptual magnet” effectively warps sensory-feature space to “pull” neighboring sounds towards a mental reference point (Kuhl, 1991; Kuhl, 2004). Evidence from human infants (Eimas et al., 1971; Kuhl, 1991; Rosen & Howell, 1987) and animals (e.g., Kuhl & Miller, 1975) lends support for an innate (or at least ontogenetically early) basis of categorization that arises within the auditory system.

However, some researchers claim that speech processing is far more graded than original theories of categorical perception would suggest (Massaro, 1987; Toscano et al., 2010). Early definitions of categorical perception assume that irrelevant, fine-grained phonetic details are discarded in favor of higher-level categorical representations, thus limiting discriminability between stimuli of the same category identity (e.g., Liberman et al., 1957). Gradient perception, in contrast, was thought to reflect impairments of phonological representations such as those observed in children with dyslexia or specific language impairment (Joanisse et al., 2000; McMurray et al., 2010; Werker & Tees, 1987). Yet, neurotypical, normal hearing listeners are still sensitive to within-category differences. The fuzzy logical model of perception (FLMP) asserts speech is initially processed as graded, continuous sensory information then classified at a higher-level, decisional stage according to relative goodness-of-fit to categorical representations (Massaro, 1987; Massaro & Oden, 1980; Oden & Massaro, 1978). Graded representations of speech stimuli may actually be more advantageous to listeners for resolving perceptual ambiguities or facilitating lexical access (Clayards et al., 2008; Kapnoula et al., in press; Kapnoula et al., 2017; Kong & Edwards, 2016; McMurray et al., 2002, 2009). Some even argue that categorical perception is entirely an artifact of the decisional processes required for binary identification and forced-choice discrimination tasks common to these experiments (Crowder, 1989; Gerrits & Schouten, 2004; Massaro, 1987; Schouten et al., 2003; Toscano et al., 2010). Additionally, Schouten et al. (2003) demonstrated that discrimination of speech is merely a relative comparison of acoustic features and does not depend on phoneme classification. Taking this perspective into consideration, we prefer a more general approach to categorical perception throughout this dissertation, one that focuses on categorization as a classification and labeling process (Medin & Barsalou, 1987).

Many of these early theories were unclear on how *learning* and experience played a role in auditory categorization. Lane (1965) was one of the first to propose that categories are not predetermined by either articulatory or auditory constraints but rather from learning the categorical labels for auditory sounds. He showed that brief categorization training of nonspeech stimuli was sufficient to induce sharp identification and discrimination functions indicative of categorical perception. Additional research has demonstrated that categorical boundaries are flexible and can be shaped by experiential factors such as native language exposure in infancy (Kuhl et al., 2006; Kuhl, 1991), knowledge of a mental lexicon (e.g., Ganong effect; Ganong, 1980; Myers & Blumstein, 2008), and long-term auditory experience such as music training (Bidelman et al., 2014; Burns & Ward, 1978; Howard et al., 1992; Siegel & Siegel, 1977; Zatorre & Halpern, 1979). These results suggest that categorical perception is more than simply a natural or innate property of the auditory system; rather, auditory categories can be learned and influenced by both short- (e.g., training) and long-term auditory experience (e.g., formal music practice) (Rosen & Howell, 1987).

Computational modeling approaches have offered additional insight into the evolving nature of categorical perception as the models must “learn” to uncover the underlying structure for accurate categorization. For example, the TRACE model of speech perception assumes that excitatory and inhibitory feedback connections refine target feature or phoneme inputs from higher, more abstract processing levels (e.g., words) on-line during perception (McClelland & Elman, 1986). Alternatively, some models facilitate category learning through backpropagation feedback (Damper & Harnad, 2000; Harnad et al., 1991; Rumelhart et al., 1986), competitive activations among category-detecting neurons (Goldstone et al., 1996), or modifying activation weights of exemplar nodes during training (e.g., ALCOVE; Kruschke, 1992; Kruschke, 1993). In such cases, multiple iterations (i.e., experience or training) of stimulus input prompt adjustment of the relative connection strengths between nodes in the model to reduce output classification error. While computational models may mirror human and animal data and propose mechanistic accounts of category learning, they do not directly specify the neural mechanisms that support auditory categorization.

Neurobiology of auditory categorization

In recent decades, neuroscience research has attempted to uncover where categories reside in the brain. Electrophysiologic responses along the auditory hierarchy (measured by EEG) suggest that auditory categorization emerges by 150-200 ms post-stimulus onset, in the timeframe of the P2 event-related potential (ERP) component generated primarily within auditory cortex (Alho et al., 2016; Bidelman & Alain, 2015; Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017, 2019; Bidelman et al., 2014; Crowley & Colrain, 2004; Ross et al., 2013; Toscano et al., 2018). Short-term categorization training induces changes in sensory encoding processes (Liu & Holt, 2011; Reetzke et al., 2018; Ross et al., 2013; Tremblay et al., 2001; Tremblay et al., 2009). Specifically, identification training leads to a decrease in perceptual sensitivity in auditory cortex for within-category stimuli while learning to discriminate sounds promotes enhanced sensitivity for differences between stimuli (Guenther et al., 2004). Long-term experiences such as music training (Bidelman & Alain, 2015; Bidelman & Walker, 2017, 2019; Bidelman et al., 2014; Wu et al., 2015) or language expertise (Bidelman & Lee, 2015; Zhang et al., 2005) have been associated with faster, more robust sensory-ERP responses of categorical stimuli. Collectively, these results suggest that categorical representations arise in low-level, perceptual processes along the auditory pathway and are tuned through experience.

Source analysis and functional magnetic resonance imaging (fMRI) techniques indicate a distributed frontotemporal brain network primarily in the left hemisphere for speech categorization (Alho et al., 2016; Bidelman & Walker, 2019; Binder et al., 2004; Bouton et al., 2018; Chang et al., 2010; Feng et al., 2017; Golestani & Zatorre, 2004; Lee et al., 2012; Liebenthal et al., 2010; Myers et al., 2009; Myers & Swan, 2012). Sensitivity to graded acoustic-phonetic features is often observed in left superior temporal gyrus (STG) while phoneme category selectivity engages primarily left inferior frontal gyrus (IFG; Alho et al., 2016; Myers et al., 2009; Toscano et al., 2018). A few studies have reported categorical effects for speech in the temporal lobe as well (Chang et al., 2010; Desai et al., 2008; Joanisse et al., 2007; Liebenthal et al., 2005). Evidence for auditory categories in the left hemisphere is mainly limited to the speech domain. In contrast, a few studies with trained musicians suggest parallel neural circuitry in the right hemisphere for music interval and chord categorization (Bidelman & Walker, 2019; Klein & Zatorre, 2011, 2015).

Some studies suggest that category learning requires mainly higher-level, frontal brain areas (e.g., left IFG, cingulate; Luthra et al., 2019; Myers & Swan, 2012) while others claim category formation occurs in temporal regions (e.g., STG, middle temporal gyrus; Desai et al., 2008; Liebenthal et al., 2010). Rapid changes in both temporal and frontal source-localized activity have also been observed following short-term frequency discrimination training (de Souza et al., 2013) and tone language learning (Lee et al., 2017). Additionally, successful speech categorization (Fuhrmeister & Myers, 2021) and nonnative phoneme learning (Golestani et al., 2007; Wong et al., 2008) have been attributed to preexisting structural differences within the auditory cortex. Despite the variety of brain regions that are sensitive to category-level information, it is clear both sensory (e.g., STG) and cognitive regions (e.g., IFG) play critical roles in categorical processing and acquisition of novel categories.

Individual differences in auditory categorization

Individuals naturally vary in their abilities to categorize auditory sounds (e.g., Hazan & Rosen, 1991; Howard et al., 1992; Kong & Edwards, 2011, 2016; Silva et al., 2020). Much of the research surrounding individual differences in categorical perception has focused on cue-weighting strategies in speech perception (e.g., Hazan & Rosen, 1991; Idemaru et al., 2012; Kapnoula et al., 2017; Kong & Edwards, 2011). For example, listeners that display more gradient responses for speech identification also exhibit greater sensitivity to a secondary acoustic cue, fundamental frequency, suggesting that variability in speech perception is consistent and systematic within individuals (Kong & Edwards, 2011, 2016). These differences cannot be explained by general cognitive control or executive functions (Kapnoula et al., 2017; Kong & Edwards, 2016).

The neural underpinnings of speech perception and categorization variability are not well understood. We have previously shown that individuals with superior (music) perceptual listening abilities exhibit enhanced brainstem encoding for speech embedded in a noisy background (Mankel & Bidelman, 2018). Machine learning techniques have linked speech identification reaction times with functional connectivity differences between several brain regions, particularly right paracentral and left middle temporal gyri (Al-Fahad et al., 2020). More successful speech category learners generally show greater neural activation after training, particularly in the STG (Díaz et al., 2008; Kajiura et al., 2021; Wong et al., 2007). Structural differences within primary auditory cortex have also been associated with enhanced categorization of both native speech (Fuhrmeister & Myers, 2021) and learned, nonnative phonemes (Golestani et al., 2007; Golestani et al., 2002; Wong et al., 2008). These studies suggest that individual differences in categorical perception my at least be partially driven by sensory encoding processes within auditory cortex, but further work is needed to better characterize the underlying neurobiology of individual categorization and learning performance.

Present work

This dissertation was motivated to answer the following question: what makes someone successful in identifying and categorizing auditory sounds? Specifically, it is unclear whether variability in categorization performance—beyond long-term auditory experience or cue-weighting strategies—is influenced by differential neural encoding processes in the brain. Cross-sectional studies in highly trained listeners (e.g., musicians) do not account for innate advantages in auditory function that might enhance auditory categorization. Our prior work revealed an interplay between preexisting and experience-driven neural factors in speech perception. While a “musician’s advantage” for speech processing has been previously reported (e.g., Başkent &

Gaudrain, 2016; Besson et al., 2011; Parbery-Clark et al., 2009; Slater et al., 2015; Zendel et al., 2015), our study demonstrated that even non-musicians with highly adept musical processing abilities ("musical sleepers"; Law & Zentner, 2012) exhibit enhanced neural encoding for speech-in-noise (Mankel & Bidelman, 2018). Chapter 2 extended Mankel and Bidelman (2018) by evaluating whether these differences in musical listening skills are associated with speech identification abilities. Specifically, this study used EEG to assess whether inherent musicality predicts speech categorization (i.e., vowels) at both behavioral and neural levels. We hypothesized that "musical sleepers" would show superior speech identification and enhanced auditory-evoked neural responses compared to less adept listeners.

Another open question is whether these neural processing differences offer perceptual advantages in learning new auditory categories. Given its implications for second language acquisition, a significant portion of research on the neurophysiology of successful auditory category learning has focused on non-native phoneme identification (e.g., Díaz et al., 2008; Kajiura et al., 2021; Myers & Swan, 2012; Wong et al., 2007). Although evidence suggests parallel neural changes following speech and nonspeech training (Liu & Holt, 2011), how nonspeech categories develop in the brain through learning is not well understood. Chapter 3 investigated the neurobiological mechanisms that contribute to the successful acquisition of nonspeech (i.e., music) auditory categories. Listeners were rapidly trained to identify musical intervals (~20 minutes of training). Individual categorization performance was evaluated against neural measures—both preexisting structural differences (MRI) as well as functional, neuroplastic changes (EEG) that arise after training—to characterize the nature/nurture processes in the brain that contribute to successful auditory category learning. Given evidence for left versus right hemispheric asymmetries in the neural processing of speech and music categories, respectively (Bidelman & Walker, 2019; Binder et al., 2004; Bouton et al., 2018; Klein & Zatorre, 2011, 2015; Lee et al., 2012; Myers et al., 2009), we speculated that the successful acquisition of musical interval categories would involve a layering of both structural and functional differences, particularly in right auditory cortex.

Chapter 2

Auditory categorical processing for speech is modulated by inherent musical listening skills

Introduction

Music training is associated with enhanced auditory processing, including the categorical perception (CP) of speech (Bidelman & Alain, 2015; Bidelman & Walker, 2019; Bidelman et al., 2014; Wu et al., 2015) and music stimuli (Howard et al., 1992). Musicianship has been attributed to sharper speech identification (Bidelman & Walker, 2019; Bidelman et al., 2014), faster labeling speeds (Bidelman & Alain, 2015; Bidelman et al., 2014), and higher perceptual sensitivities for within- vs. between-category contrasts (Wu et al., 2015). Whether these speech advantages truly reflect experience (training) or other predispositions remains unclear. Musicians' improved speech processing may instead reflect increased motivation (McAuley et al., 2012), enhanced auditory attention and cognitive advantages (Yoo & Bidelman, 2019), or even innately superior auditory abilities (e.g., musicality) that exist without formal music training (Mankel & Bidelman, 2018; Swaminathan & Schellenberg, 2017).

While musicality is a multidimensional trait, we focus on abilities pertaining to complex listening, following a long tradition of describing musical aptitude via receptive skills (Gordon, 1965; Law & Zentner, 2012; Mankel & Bidelman, 2018). Recently, we demonstrated even nonmusicians vary in innate music abilities and people with exceptionally high levels of musicality show stronger neural processing of speech (Mankel & Bidelman, 2018), a finding typically attributed to trained musicians (Bidelman et al., 2014; Parbery-Clark et al., 2009). Positive associations between phoneme discrimination and musicality (but not music training itself) have been observed even after controlling for cognitive and socioeconomic status (SES) (Swaminathan & Schellenberg, 2017). Irrespective of formal training, music aptitude also predicts stronger CP for musical chords (Howard et al., 1992). Such findings highlight the need to distinguish innate from experience-dependent characteristics for defining neuroplasticity. Presumably, individual differences in auditory processing may at least partially drive sound categorization skills, which would temper assumptions that experience, *per se*, drives musicians' enhancements in speech processing (Mankel & Bidelman, 2018; Swaminathan & Schellenberg, 2017).

Neural evidence suggests CP emerges in the brain by 100-200 ms, in the timeframe of the N1-P2 event-related potentials (ERPs; Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2019; Dittinger et al., 2018). Enhanced N1-P2 amplitudes, for example, have been linked with stronger identification (Bidelman & Walker, 2017) and faster labeling speeds during speech categorization (Bidelman & Alain, 2015; Bidelman et al., 2014). Categorical effects have also been reported for the later N2 and P3 waves, whose amplitudes differentiate stimuli at the categorical boundary and between native vs. nonnative phonetic contrasts (Dittinger et al., 2018). Neuroimaging studies have further implicated several brain regions in auditory categorical processing, including the superior temporal gyrus/sulcus, middle temporal gyrus, premotor cortex, supramarginal gyrus, inferior parietal cortex, planum temporale, and inferior frontal gyrus (Bidelman & Walker, 2019; Guenther et al., 2004). These neural signatures of CP are also associated with formal music training (Bidelman & Alain, 2015; Bidelman & Walker, 2019; Bidelman et al., 2014; Dittinger et al., 2018; Elmer et al., 2012). While long-term auditory experience might tune categorical speech processing, it is unclear (i) how much CP is susceptible to individual differences and (ii) whether musicians' enhancements in this process might be partially due to innate auditory sensitivities irrespective of formal experience.

Here, we tested whether individual differences in musicality—in the absence of formal music training—affect speech categorization. To this end, we measured EEGs in nonmusicians who varied in their musical listening abilities as they classified vowels along an acoustic-phonetic continuum. If individual differences in auditory perceptual skills drive CP, we predicted that individuals with higher levels of innate musicality ("musical sleepers"; Law & Zentner, 2012; Mankel & Bidelman, 2018) would show stronger behavioral CP and neural responses to speech. Our data confirm that musical sleepers categorize speech more efficiently, as evidenced by stronger categorical coding compared to those with poorer music aptitude.

Methods

Participants

The sample included N =14 young adults (7 females; age: $\mu \pm \sigma = 24.9 \pm 1.7$ years). This sample size is comparable to studies assessing musicians and nonmusicians (Bidelman & Walker, 2019; Bidelman et al., 2014) and allowed us to test the premise that such variations in CP might result from differences in inherent musicality rather than formal training. Participants were right-handed, had normal hearing (thresholds ≤ 25 dB HL, 250-8000 Hz), no tone language experience or neurological disorders, and a collegiate level of education (17.3 ± 3.00 years). Average parental education, a common measure of SES [highest parental education: 1 (high school without diploma or GED) to 6 (doctoral degree)], was 4.14 ± 0.57 indicating a bachelor's level education (Entwislea & Astone, 1994). To isolate effects of inherent music listening abilities on auditory processing (Mankel & Bidelman, 2018), participants were required to have < 3 years total of formal music training (0.57 ± 0.76 years) and no music experience within the past 5 years. All gave written consent according to a protocol approved by the UofM Institutional Review Board.

Behavioral test of musicality

The brief Profile of Music Perception Skills (PROMS) assessed aptitude related to receptive musical skills (Law & Zentner, 2012). This test comprises same-different tasks of melody, tuning, accent, and tempo discrimination. Each subtest contains 18 trials, and higher scores indicate stronger musical listening abilities (max score=72). The PROMS can differentiate listeners based on their musical experience (professional vs. amateur musicians vs. nonmusicians), and is sensitive to detect untapped musical potential among nonmusicians ("musical sleepers"; Law & Zentner, 2012; Mankel & Bidelman, 2018). For details see (Law & Zentner, 2012).[1]

Stimuli

We used a synthetic five-step vowel continuum to assess CP for speech (see Fig. 1 in ref. Bidelman et al., 2014). Tokens were 100 ms. Fundamental (F0), second (F2), and third formant (F3) frequencies were identical across tokens (F0: 100 Hz; F2: 1090 Hz; and F3: 2350 Hz). First formant (F1) was varied across five equidistant steps (430 to 730 Hz), yielding a perceptual continuum from /u/ to /a/.

Behavioral data

The task was identical to our recent reports (Bidelman, Moreno, et al., 2013; Bidelman et al., 2014). Stimuli were delivered binaurally through ER-2 earphones at 82 dB SPL. Listeners heard 200 trials of each speech token (random order) and were instructed to label them as “u” or

---

[1] Both the brief and full PROMS have been cross validated in large cohort studies with trained musicians as well as other established tests of musical abilities (Kunert et al., 2016; Law & Zentner, 2012). PROMS scores show discriminant validity from basic psychoacoustic skills (temporal gap detection; Law & Zentner, 2012) as well as cognitive functions such as working memory (Kunert et al., 2016), thus motivating its use here as a measure of preexisting, receptive musical abilities in nonmusicians.

“a” via the keyboard as fast and accurately as possible. The interstimulus interval was 400-600 ms (20 ms steps; jittered).

Individual identification curves were fit with a sigmoid: $P = 1/[1 + e^{-\beta 1(x-\beta 0)}]$, where P is the proportion of trials labeled as /a/, x is the token number, and β0 and β1 are the location and slope parameters of the psychometric fit. Large β1 values indicate steeper slopes and stronger CP. Response times (RTs) were computed as the average labeling speed per token. RTs outside 250-2500 ms were excluded as fast guesses or lapses in attention (Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017).

EEG recordings

EEGs were recorded from 64 electrodes at 10-10 scalp locations (500 Hz sample rate) (Synamps RT amplifiers, Neuroscan). Impedances were <10 kΩ. Pre-processing was performed in BESA® Research (v7) (BESA, GmbH). Blinks were nullified in the continuous data via spatial filtering (Gordon, 1965). Trials with voltages >±120 µV were discarded from averaging. Recordings were epoched (-200-800 ms), baseline corrected, filtered from 1-30 Hz, and averaged across trials to compute ERPs for each speech token per listener.

Source analysis

We transformed raw ERPs to source space using BESA’s AEP virtual source montage (Scherg & von Cramon, 1986). We then extracted source waveforms (units nAm) from the radially-oriented dipoles in left and right auditory cortices (AC) to assess categorical speech coding (Talairach coordinates of [-37, -18, 17] and [37, -18, 17], respectively) (e.g., Bidelman & Lee, 2015; Bidelman & Walker, 2019). Neural correlates of CP emerge around the N1 and P2 waves (Bidelman & Alain, 2015; Bidelman & Lee, 2015; Bidelman, Moreno, et al., 2013). Thus, N1 was measured as the peak negativity between 100-160 ms; P2 as the peak positivity between

160-220 ms. Categorical coding in the ERPs was evaluated as the difference between P2 responses to prototypical tokens (endpoints) and the ambiguous midpoint [i.e., mean (Tk1, Tk5) - Tk3; ΔP2] (Bidelman & Walker, 2017, 2019); more negative ΔP2 latencies mirror categorical processing whereby /u/ and /a/ exemplars elicit faster latencies than stimuli at the ambiguous midpoint of the continuum (i.e., Tk1/5 < Tk3).

Results

Behavioral data

Listeners were highly separable based on a median split of their PROMS scores, revealing some nonmusicians have inherently better music listening skills (**Fig. 1A**) (Mankel & Bidelman, 2018). Low and high musicality groups did not differ in age, handedness, years of musical training, years of education, or SES (all *P*s > 0.05). The high PROMS group demonstrated superior performance not only on the PROMS total score (Wilcoxon; normal approximation, one sided: $Z = 1.79$, $P = 0.001$) but also the tempo ($Z = 1.79$, $P = 0.0363$) and tuning subtests ($Z = 2.78$, $P = 0.0027$).

Groups did not differ in their slopes of speech identification ($Z = -0.9594$, $P = 0.3374$; **Fig. 1B**) but the high PROMS group showed marginally faster RTs across the continuum ($F_{1,52} = 3.87$, $P = 0.0544$, Cohen's $d = 1.136$; **Fig. 1C**). Higher PROMS scores were also associated with faster RTs ($r = -0.55$, $P = 0.0435$; **Fig. 1D**). Thus, while speech CP was not stronger in one group over the other, *per se*, faster RTs suggest individuals with better music listening skills were more efficient at making speech categorization judgments.

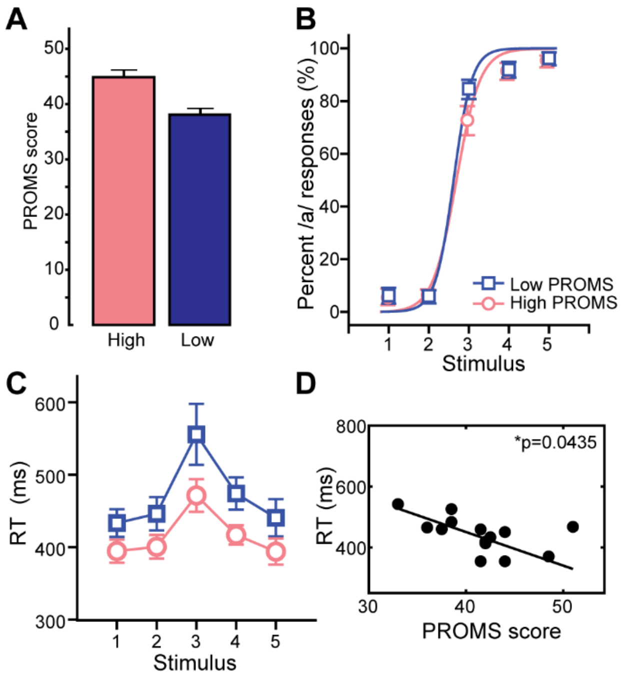


Figure 1: Behavioral data. (**A**) PROMS scores reveal some individuals have naturally more adept auditory perceptual skills in the absence of any formal music training (median split of into high- and low-musicality groups). High and low PROMS groups did not differ in identification (**B**), but highly musical listeners were more efficient (faster) at identifying categorical speech stimuli (**C**). (**D**) Better musical perceptual skills correlated with faster average RTs. *$P < 0.05$, error bars ± 1 SEM. PROMS, Profile of Music Perception Skills.

Source event-related potentials

Fig. 2 shows scalp topographies and source waveforms from left/right AC. Two-way, mixed model ANOVAs evaluated group x hemisphere effects on ERP measures (random factor = subjects nested within group). Pooled across tokens, we found sole hemispheric difference in N1 latencies (LH > RH; $F_{1,128} = 7.15$, $P = 0.0085$, $d = 1.54$). Subsequent analysis focused on changes in P2 (i.e., ΔP2) between phonetic (Tk1/5) and non-phonetic (Tk3) tokens as a marker of speech identification and categorical processing (Bidelman & Alain, 2015; Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017, 2019; Bidelman et al., 2014).

ΔP2 latencies showed a group × hemisphere interaction ($F_{1,12} = 6.79$, $P = 0.023$, $d = 1.50$; **Fig. 2C**). Tukey-Kramer contrasts revealed the high PROMS group had more negative (i.e., more categorical) ΔP2 latencies in the RH than the low PROMS group ($t_{12} = -2.33$, $P = 0.038$).

Stronger speech differentiation, as indexed by ΔP2 latencies, was also observed in RH vs. LH specifically in the high PROMS group (RH < LH; $t_{12} = 3.01$, $P = 0.0108$). No other contrasts including amplitude measures were significant.

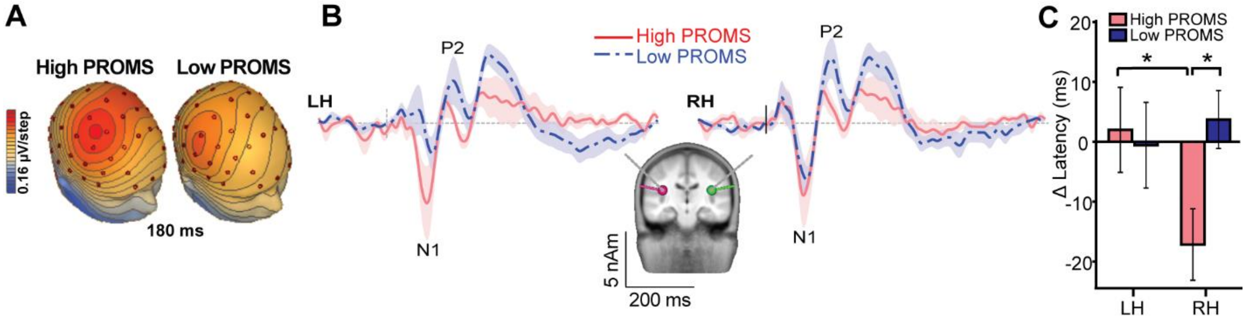


Figure 2: Neural data. ERPs reflect categorical coding around the timeframe of P2. (**A**) Scalp topographies (180 ms) and (**B**) grand average auditory cortical source waveforms by PROMS groups. (**C**) Phonetic (Tk1/5) vs. non-phonetic (Tk3) contrast of P2 (i.e., ΔP2) assessed categorical processing, which was stronger for individuals with better musical listening abilities. ΔP2 latencies were more categorical (i.e., negative) in the right hemisphere for the high PROMS group. $^*P < 0.05$, error bars/shading ± 1 SEM.

Discussion

Our findings show "musical sleepers" who exhibit naturally higher levels of musicality (but are nevertheless nonmusicians) are more efficient at categorizing speech sounds (i.e., faster RTs). Neural responses mimicked behavioral benefits in that high PROMS listeners showed more categorical ERPs within right AC. Our results provide novel evidence that listeners' efficiency of categorical processing at both behavioral and neural levels varies with their inherent auditory perceptual skills.

Critically, our data reveal natural auditory sensitivities, in the absence of music training, are associated with improvements in auditory categorical processing. While both groups demonstrated categorical speech processing—sharp transitions at the perceptual boundary and delayed RTs for ambiguous relative to prototypical tokens (Bidelman et al., 2019; Pisoni & Tash, 1974)—those with higher inherent musicality exhibited faster RTs. Interestingly, the faster RTs observed in our high music aptitude listeners mirrors results in formally trained musicians, who also show faster speech categorization (Bidelman & Alain, 2015; Bidelman et al., 2014; Dittinger et al., 2018). While faster RTs in the high PROMS group may reflect faster initiation/execution of motor responses, group differences in neural activity occurred as early as P2 (i.e., well before RTs). This implies pre-cognitive, pre-motor brain activity (i.e., sensory coding) contributes to improved categorization efficiency in high aptitude nonmusicians.

That RTs but not identification differed between groups suggests inherent auditory skills improve the decision processes and/or speed of access to internalized categories rather than a sharpening of those phonetic representations, *per se* (Bidelman et al., 2019; Pisoni & Tash, 1974). A lack of difference in identification may have been expected given that our listeners were native speakers, highly familiar with English vowel contrasts (Bidelman & Alain, 2015; Dittinger et al., 2018; Foster & Zatorre, 2010b). In comparison, categorical identification (in addition to RTs) is more acute in actual *trained* musicians (cf. Bidelman & Walker, 2019; Bidelman et al., 2014). Collectively, this and other studies lead us to infer that long-term music experience improves sound classification accuracy and speeded access to speech representations above and beyond natural or inherent abilities observed here.

At the neural level, we found stronger categorical neural encoding of speech in people with better music perceptual skills. More categorical responses were observed in the right AC of the high compared to the low PROMS group. Our data here (all nonmusicians) parallel recent findings in highly experienced musicians which report plasticity in AC activity is associated with more categorical coding of sound (Bidelman & Walker, 2019). Previously, we asserted musicality-related differences in auditory CP require (i) strong experiential plasticity rather than subtle innate function, or (ii) tasks requiring top-down processing and/or attention (Mankel &

Bidelman, 2018). Results of this study extend previous findings by demonstrating both behavioral and ERP enhancements in musical sleepers during active speech identification. Results here show that certain individuals exhibit naturally superior auditory system function, specifically in the right AC and timeframe of P2 (~160-200 ms), which contributes to enhanced speech perception.

Given the known LH dominance of language processing, it is surprising that categorial neural responses to speech were more apparent in *right* AC. One possibility is that highly musical individuals recruit additional resources in RH for speech processing, which could contribute to faster ERP latencies and RTs seen here. Indeed, greater training-related plasticity is observed in right compared to left AC (Bermudez et al., 2009). However, previous cross-sectional studies do not isolate innate differences of brain structure and function prior to training. Although the role of "native abilities" have been acknowledged in studies on musicianship and auditory perceptual abilities (Foster & Zatorre, 2010b), they have never been fully tested to the extent here (cf. Mankel & Bidelman, 2018).

While we do not refute the existence of experience-dependent plasticity of musicianship (Habib et al., 2016; Mankel & Bidelman, 2018), the "musical brain" is likely an interplay between predispositions, environmental factors, and training (Mankel & Bidelman, 2018). Our data reveal inherent differences in brain function can contribute to more efficient and robust categorical speech processing. Thus, inherent auditory skills may at least partially contribute to experience-dependent neuroplastic effects reported in studies on musical training and speech-language function (Bidelman & Walker, 2019; Bidelman et al., 2014). Other studies have similarly shown associations between speech discrimination and musicality (i.e., rhythm perception) but not music training itself (Swaminathan & Schellenberg, 2017). While music-based interventions for communication disorders are promising (Habib et al., 2016), care should be taken to understand individuals’ experience and inherent auditory function in order to maximize learning or therapeutic benefits. We add to the growing body of evidence that some individuals have naturally superior auditory abilities which enables better neural processing and perception of speech (Mankel & Bidelman, 2018; Swaminathan & Schellenberg, 2017).

Chapter 3

Functional plasticity coupled with structural predispositions in auditory cortex shape successful music category learning

Introduction

Classifying continuously varying sounds into meaningful categories like phonemes or musical intervals enables more efficient processing of an auditory scene (Bidelman et al., 2020). Categorization of auditory stimuli is also a foundational skill for language development and is believed to arise from both learned and innate factors (Livingston et al., 1998; Mankel, Barber, et al., 2020; Mankel, Pavlik Jr, et al., 2020; Perez-Gay Juarez et al., 2019; Rosen & Howell, 1987). Auditory categories are further shaped by experiences such as speaking a second language (Escudero et al., 2011; Lively et al., 1993; Perrachione et al., 2011) or musical training (Bidelman & Walker, 2019; Bidelman et al., 2014; Wu et al., 2015), suggesting flexibility in categorical perception with learning. While the behavioral aspects of category acquisition are well documented, the underlying neural mechanisms and the influence of individual differences in shaping this process are poorly understood.

Characterizing the neurobiology of category acquisition is typically confounded by prior language experience and the overlearned nature of speech (Liu & Holt, 2011). For example, perceptual interference from native-language categories can impede the learning of foreign speech sounds (Flege & MacKay, 2004; Francis et al., 2008; Guion et al., 2000). Instead, nonspeech stimuli (e.g., music) offers the ability to probe the neural mechanisms of category learning without the potential confounds of language background or automaticity that stems from using speech materials (Goudbeek et al., 2009; Guenther et al., 1999; Liu & Holt, 2011; Smits et al., 2006; Yi & Chandrasekaran, 2016). In this regard, musical categories (i.e., intervals, chords) offer a fresh window into tabula rasa category acquisition. Indeed, nonmusicians are unable to adequately categorize musical stimuli despite their exposure to music in daily life (Bidelman & Walker, 2019; Howard et al., 1992; Klein & Zatorre, 2011; Locke & Kellar, 1973; Siegel & Siegel, 1977). While several studies have assessed category learning of musical intervals, they either used highly trained listeners (Burns & Ward, 1978) or focused on different training methods that maximize learning gains (Little et al., 2019; Pavlik Jr et al., 2013). To our knowledge, no study has assessed the *neural* changes associated with category learning in music.

Speech categorization is believed to emerge in the brain by around N1 of the cortical event-related potentials (ERPs) and fully manifests by P2 (i.e., ~150-200 ms; Alho et al., 2016; Bidelman & Lee, 2015; Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017; Mankel, Barber, et al., 2020; Ross et al., 2013). Fewer studies have examined the electrophysiological underpinnings of music categorization (i.e., intervals, chords), but evidence from musicians suggests a similar neural time course (Bidelman & Walker, 2019). Functional magnetic resonance imaging suggests that categorization training leads to a decrease in perceptual sensitivity for within-category stimuli in auditory cortex while learning to discriminate categorical sounds shows the opposite effect—greater sensitivity to differences between stimuli (Guenther et al., 2004). Still, the majority of studies on category learning have involved speech. Although there are probably some parallels (Liu & Holt, 2011), it remains unclear whether the neuroplastic changes that arise when rapidly learning nonspeech categories (e.g., music) parallels that of speech.

More generally, auditory perceptual learning studies have reported changes in both early sensory-evoked (i.e., N1, P2) and late slow-wave ERP responses following training (Alain et al., 2010; Alain et al., 2007; Atienza et al., 2002; Ben-David et al., 2011; Bosnyak et al., 2004; Carcagno & Plack, 2011; Tong et al., 2009; Tremblay et al., 2001; Tremblay & Kraus, 2002;

Tremblay et al., 2009; Wisniewski et al., 2020). A true biomarker of learning, however, should vary with learning performance (Tremblay et al., 2014). Because modulations in P2 amplitudes can occur with mere passive stimulus exposure in the absence of training improvements, some posit P2 reflects aspects of the task acquisition process rather than training or perceptual learning, *per se* (Ross et al., 2013; Ross & Tremblay, 2009; Tremblay et al., 2014). Given the equivocal role of P2 in relation to auditory learning, we aimed to clarify whether changes in sensory-evoked responses (i.e., P2) and/or late slow-wave neural activity scale with individual behavioral outcomes during the rapid learning of novel music categories.

There is also significant variability in the acquisition of auditory categories (e.g., Golestani & Zatorre, 2009; Howard et al., 1992; Mankel, Pavlik Jr, et al., 2020; Silva et al., 2020), especially for speech (Díaz et al., 2008; Fuhrmeister & Myers, 2021; Kajiura et al., 2021; Mankel, Barber, et al., 2020; Wong et al., 2007). More successful learners show greater neural activation, particularly in auditory cortex (Díaz et al., 2008; Kajiura et al., 2021; Wong et al., 2007). Such variability might be attributable to differences in the creation or retrieval of long-term memories for prototypical vs. non-prototypical sounds during learning (Golestani & Zatorre, 2009). However, we have previously demonstrated better categorizers show greater efficiency in early sensory encoding processes (~150-200 ms), suggesting stimulus representations themselves are tuned at the individual level rather than later memory-related processes, *per se* (Mankel, Barber, et al., 2020).

In addition to differences in functional processing, individual categorization abilities may be partially driven by preexisting structural advantages within the brain (Fuhrmeister & Myers, 2021). Paralleling the left hemisphere bias for speech (Binder et al., 2004; Bouton et al., 2018; Lee et al., 2012; Myers et al., 2009), categorization of musical sounds is believed to involve a frontotemporal network in the right hemisphere, including key brain regions such as the primary auditory cortex (PAC), superior temporal gyrus (STG), and inferior frontal gyrus (IFG) (Bidelman & Walker, 2019; Klein & Zatorre, 2011, 2015; Mankel, Barber, et al., 2020). PAC/STG size (primarily right hemisphere) has also been associated with perception of relative pitch and musical transformation judgments (Foster & Zatorre, 2010a), melodic interval perception (Li et al., 2014), spectral processing (Schneider et al., 2005), and even musical aptitude (Schneider et al., 2002). To our knowledge, few studies have examined the structural correlates of categorization differences on the individual level. In the domain of speech, faster, more successful learners of nonnative phonemes exhibit larger left Heschl's gyrus (Golestani et al., 2007; Wong et al., 2008) and parietal lobe volumes (Golestani et al., 2002). Additionally, better and more consistent speech categorizers show increased right middle frontal gyrus surface area and reduced gyrification in bilateral temporal cortex (Fuhrmeister & Myers, 2021). We thus hypothesized that successful category learning in *music* would be predicted by neuroanatomical differences (e.g., gray matter volume, cortical thickness), with perhaps effects favoring right PAC.

The aim of this study was to examine the functional and structural neural correlates of auditory category learning following short-term identification training of music sound categories. Musical intervals allowed us to track sound-to-label learning without the potential lexical-semantic confounds inherent to using speech materials (Liu & Holt, 2011). We measured learning-related changes in the cortical ERPs in musically naïve listeners against a control group that did not receive identification training to determine the specificity of neuroplastic effects. If rapid auditory category learning is related to enhanced sensory encoding of sound, we predicted changes in early brain activity manifesting at or before auditory object formation (i.e., prior to

~250 ms; P2). If instead, short-term learning is associated with later cognitive processes related to decision and/or task strategy, we expected neural effects to emerge later in the ERP time course (e.g., late slow waves > 400-500 ms; Alain et al., 2007). Additionally, we anticipated successful learners would recruit neural resources in right auditory cortices, mirroring the left hemispheric specialization supporting speech categorization (Bidelman & Walker, 2019; Joanisse et al., 2007; Klein & Zatorre, 2011; Liebenthal et al., 2005). Our findings show that successful auditory category learning is characterized by both structural and functional differences in right auditory cortex. The presence of anatomical differences along with ERP changes specific to learning suggest that the acquisition of auditory categories depend on a layering of preexisting and short-term plastic changes in brain function.

Materials & Methods

Participants

Our sample included N=33 participants. Nineteen young adults (16 females) participated in the learning group with behavioral and neural measurements taken at pre-test, during active identification training, and post-test phases. An additional fourteen (7 females) served as a no-contact control group (previously collected unpublished data; see Mankel, Barber, et al., 2020) and only participated in the pre- and post-test blocks (i.e., they did no go through training). These group sizes are comparable to other studies investigating the neural mechanisms of category learning (e.g., Liu & Holt, 2011; Myers & Swan, 2012; Wong et al., 2007) and allow us to test whether observed neuroplastic changes are specific to learning. All had normal hearing (thresholds ≤25 dB SPL, 250-8000 Hz), were right-handed (Oldfield, 1971), and had no history of neurological disorders. Participants completed questionnaires that assessed education level, socioeconomic status (SES) (Entwislea & Astone, 1994), language history (Li et al., 2006), and music experience. Groups were comparable in age (learners: μ = 24.9 ± 4.0 yrs, controls: μ = 24.9 ± 1.7 yrs; p = 0.549), education (learners: μ = 18.5 ± 3.3 yrs, controls: μ = 17.3 ± 3.0 yrs; p = 0.321), and SES (rating scale of average parental education from 1 [some high school education] to 6 [PhD or equivalent]; learners: μ = 4.6 ± 1.3, controls: μ = 4.1 ± 0.6; p = 0.109). All were fluent in English though six reported a native language other than English. We excluded tone language speakers as these languages improve musical pitch perception (Bidelman, Hutka, et al., 2013). To ensure participants were naïve to the music-theoretic labels for pitch intervals, we required participants have no more than three years total of formal music training on any combination of instruments and none within the past five years. Critically, groups did not differ in prior music training (learners: μ = 1.1 ± 1.0 yrs, controls: μ=0.6 ± 0.8 yrs; p=0.145). All participants gave written informed consent according to protocol approved by the University of Memphis Institutional Review Board and were compensated monetarily for their time.

Stimuli

We used a five-step musical interval continuum to assess category learning of nonspeech sounds continuum (Bidelman & Walker, 2017; Mankel, Pavlik Jr, et al., 2020). Individual notes of each dyad were constructed of complex tones consisting of 10 equal amplitude harmonics added in cosine phase. Each token was 100 ms in duration with a 10 ms rise/fall time to reduce spectral splatter. The bass note was fixed at a fundamental frequency (F0) of 150 Hz while the upper note's F0 ranged from 180 to 188 Hz (2 Hz spacing between adjacent tokens). Thus, the musical interval continuum spanned a minor (token 1) to major third (token 5). The minor-major third continuum was selected because these intervals occur frequently in Western tonal music and connote typical valence of "sadness" and "happiness", respectively, and are therefore easily described to participants unfamiliar with music-theoretic labels (Bidelman & Walker, 2017).

Moreover, without training, nonmusicians perceive musical intervals in a continuous mode indicating they are initially heard non-categorically (Bidelman & Walker, 2017, 2019; Burns & Ward, 1978; Howard et al., 1992; Locke & Kellar, 1973; Siegel & Siegel, 1977; Zatorre & Halpern, 1979).

Procedure

Participants were seated comfortably in an electroacoustically shielded booth. Stimuli were presented binaurally through ER-2 insert earphones (Etymotic Research) at ~81 ± 1 dB SPL. Stimulus presentation was controlled by MATLAB routed through a TDT RP2 interface (Tucker Davis Technologies). Categorization was assessed in a pre- and post-test phase. Following brief task orientation (~2-3 exemplars), tokens of the continuum were randomly presented on each trial. Participants were instructed to label the sound they heard as either "minor" or "major" via keyboard button press as fast and accurately as possible. The interstimulus interval was 400-600 ms (jittered in 20 ms steps) following the listener's response. No feedback was provided during the pre- or post-test. Pre- and post-test procedures were consistent across both the learners and controls as both groups completed the two separate musical interval categorization assessments during a single experimental session. The learning group received additional identification training (see *Training paradigm* below) whereas individuals in the control group were offered a brief break before starting the post-test phase.

Training paradigm

Participants in the learning group underwent a 20-min identification training between the pre- and post-test phases (all performed in a single ~3 hr period). Training consisted of 500 trials, 250 presentations each of the minor and major 3rd exemplars (i.e., tokens 1 and 5), spread evenly across 10 blocks[1]. Feedback was provided to improve accuracy and efficiency of auditory category learning (Yi & Chandrasekaran, 2016). The training procedure was conducted using E-Prime 2.0 (PST, Inc.). To reduce fatigue, participants were offered a break before and after the training phase.

EEG acquisition and preprocessing

EEG data were recorded using a Synamps RT amplifier (Compumedics Neuroscan) from 64 sintered Ag/AgCl electrodes at 10-10 scalp locations and referenced online to a sensor placed ~1 cm posterior to Cz. Impedances were <10 kΩ. Recordings were digitized at a sampling rate of 500 Hz. Preprocessing was completed in BESA Research (v7.1; BESA GmbH). Continuous data were re-referenced offline to the common average reference, epoched from -200-800 ms, filtered from 1-30 Hz (4th-order Butterworth filter), baselined to the prestimulus interval, and averaged across trials to compute ERPs for each token per electrode.

MRI segmentation and volumetrics

12 out of 19 learning group participants returned on a separate day for structural MRI scanning. 3D T1-weighted anatomical volumes were acquired on a Siemens 1.5T Symphony TIM scanner (tfl3d1 GR/IR sequence; TR = 2000 ms, TE = 3.26 ms, inversion time = 900 ms, phase encoding steps = 341, flip angle = 8°, FOV = 256 x 256 acquisition matrix, 1.0 mm axial slices). Scanning was conducted at the Semmes Murphey Neurology Clinic (Memphis, TN). All MRI T1-weighted images were initially registered to MNI ICBM 152 T1 weighted atlas with 1 x 1 x 1 mm^3 isometric voxel size using affine transformation. The inverse transformation matrix was computed and applied to the brain mask in atlas space to create brain mask specific for each subject for skull removal (Evans et al., 1993). An LPBA40 T1 weighted atlas with 2 x 2 x 2 mm^3

---

[1] Two pilot subjects received 6 and 15 blocks of training, respectively, before settling on the final 10 block training regimen.

voxel size was then used to register the images and remove the cerebellum using the atlas cerebrum mask and following the same process as above (Shattuck et al., 2008). After skull removal and cerebrum extraction, an AAL3 T1 weighted atlas with 1 x 1 x 1 mm^3 voxel size that provides parcellation of a large number of brain regions was used for extracting gray matter volume in certain regions of interest (ROIs) for each participant (Rolls et al., 2020).

Data analysis

Behavioral data

Identification curves were fit with a two-parameter sigmoid function $P = 1/[1 + e-\beta1(x-\beta0)]$, where P describes the proportion of trials identified as major, x is the step number along the stimulus continuum, $\beta0$ is the locus of transition along the sigmoid (i.e., categorical boundary), and $\beta1$ is the slope of the logistic fit. Larger $\beta1$ values reflect steeper psychometric functions and therefore better musical interval categorization performance. $\beta1$ slopes were square root transformed improve normality and homogeneity of variance. Reaction times (RTs) were computed as the listeners' median response latency for the ambiguous (i.e., token 3) and prototypical tokens (i.e., mean[tokens 1 & 5]; see ERP data), after excluding outliers outside 250-2500 ms (Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017; Mankel, Barber, et al., 2020). As an index of training success, accuracy was calculated in the learning group as the average percent correct identification across all training trials.

ERP data

We analyzed a subset of electrodes from a frontocentral cluster (mean of F1, Fz, F2, FC1, FCz, FC2; see **Fig. 1** inset) where categorical effects in the auditory ERPs are most prominent at the scalp (Bidelman & Lee, 2015; Bidelman, Moreno, et al., 2013; Bidelman & Walker, 2017; Bidelman et al., 2014). Peak latencies and amplitudes were quantified for P1(40-80 ms), N1 (70-

130 ms), and P2 (140-200 ms). The mean amplitude was also measured for slow wave activity between 300-500 ms, given prior work suggesting rapid auditory learning effects in this later time frame (Alain et al., 2010; Alain et al., 2007).

We also quantified neural responses at T7 and T8 to assess hemispheric lateralization. For these analyses, we computed difference waves derived between the ambiguous and prototypical tokens (ΔERP = mean[tokens 1 & 5] – token 3) for both the pre- and post-test (Bidelman, 2015; Bidelman & Walker, 2017; Mankel, Barber, et al., 2020). Larger ΔERP values indicate stronger differentiation of category ambiguous from category prototype sounds and thus reflect the degree of "neural categorization" in each hemisphere.

MRI data

Each participant's MRI images were registered to the AAL3 atlas, ROI masks were transformed to subject space, and ROI volumes were then calculated (cm3). Atlas registration was confirmed using SPM12 toolbox in MATLAB (Penny et al., 2011). Cortical thickness was examined using a diffeomorphic registration based cortical thickness (DiReCT) measure (Das et al., 2009). We used the OASIS atlas (Marcus et al., 2009) for the computation of cortical thickness because it provides four brain segmentation priors for parcellating cerebrospinal fluid (CSF), cortical gray matter, white matter, and deep gray matter. 3D cortical thickness maps for each subject were computed based on these priors. Thickness maps were then multiplied with the AAL3 atlas (converted to subject space) to compute the cortical thickness of each brain region mapped to their corresponding labels. Finally, the mean, standard deviation, and range of the cortical thickness measurements along with the surface area and volume of the cortical regions were computed for each ROI. Volumetrics were normalized to each participant's total intracranial brain volume to control for artificial differences across individuals (e.g., head size;

Whitwell et al., 2001). To test for hemispheric differences specific to auditory neuroanatomic measures, we restricted ROI analysis to bilateral Heschl's gyrus (PAC; Brodmann 41). MRI post-processing was performed using in-house scripts coded in Python (http://www.python.org).

Statistical analysis

ERPs were analyzed using GLME mixed-effects regression models in SAS (Proc GLIMMIX; v9.4, SAS Institute, Inc.) with subjects as a random factor and fixed effects of training phase (two levels: pretest vs. posttest), token (two levels: tokens 1&5 vs. 3) and behavioral performance (identification slopes or training accuracy; continuous measures). We also included the interaction of phase and behavioral performance to investigate whether brain-behavior correspondences change after training. Similar models were used to analyze the behavioral and MRI data. Analyses on the individual groups alone included main and interaction effects of identification slopes or training accuracy (learning group only), training phase, and stimulus token. We used a backward selection procedure to remove nonsignificant variables and report final model results throughout. Post hoc multiple comparisons were corrected using Tukey adjustments. Identification function slopes were square-root transformed to reduce right skewness. Demographic variables were analyzed using Wilcoxon-Mann-Whitney and Fischer's exact tests due to non-normality. An *a priori* significance level was set at $\alpha = 0.05$. Conditional studentized residuals, Cook's D, and covariance ratios were used to identify and exclude influential outliers (West et al., 2015).

Results

Training results

Behavioral training outcomes are plotted in **Figure 3**. On average, participants in the learning group improved 10-15% in accuracy (**Fig. 3A**; $F_{9,158} = 2.05$, $p = 0.038$) and exhibited faster RTs (**Fig. 3B**; $F_{9,149} = 3.22$, $p = 0.001$) over the course of training. Training was highly effective; most individuals averaged >80-90% identification accuracy across the 10 blocks (i.e., equivalent to the approximate performance of a musician on the same task; data not shown). N=5 "nonlearners" had training accuracies that remained near chance performance; these individuals were removed for subsequent analysis. Post hoc analyses revealed RTs became faster following the second training block (all *p*'s < 0.05; block 1 vs. 2 $p = 0.052$). Similarly, listeners' identification was more accurate starting at the 9th training block compared to the first block (block 9 vs. 1: $t_{158} = 3.44$, $p = 0.025$; block 10 vs.1: $t_{158} = 3.40$, $p = 0.028$).

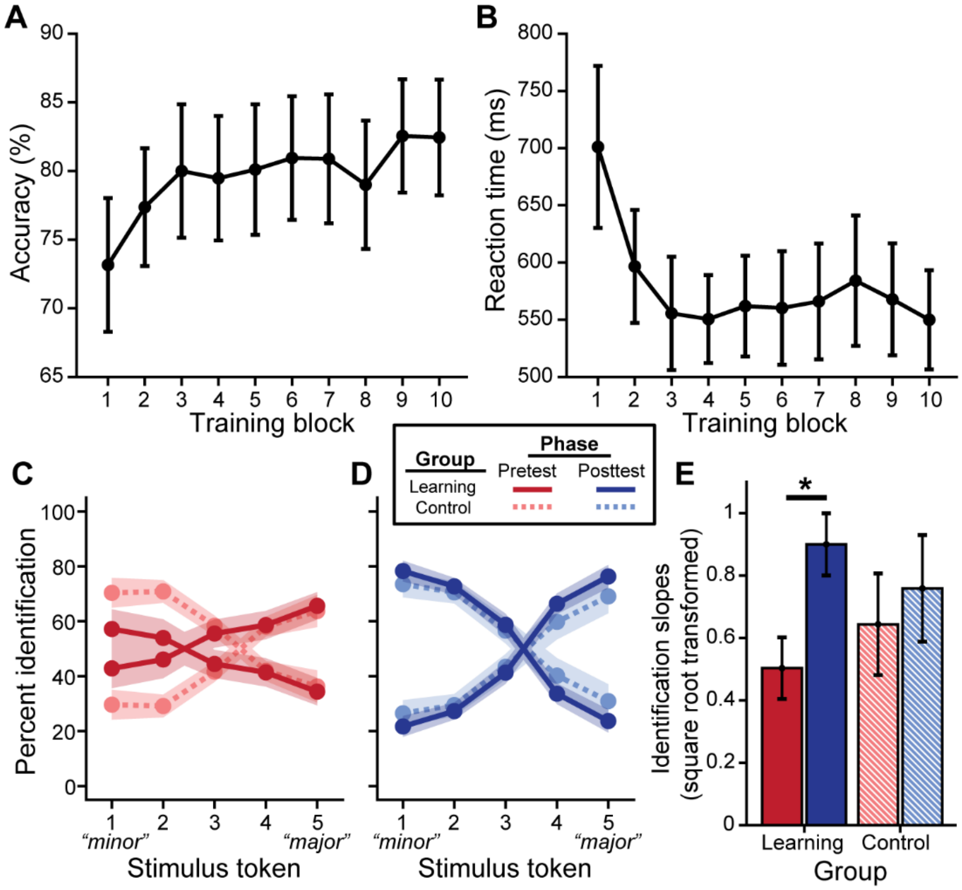


Figure 3: Behavioral categorization improves following rapid auditory training. Brief major/minor categorization training yields an increase in accuracy (A) and decrease in reaction time (B) across blocks. Pretest (C) and posttest (D) psychometric identification functions show stronger categorization for musical intervals after training for the learning group (excluding data from n=5 nonlearners); performance was identical pre- to post-test for control listeners (E). Error bars/shading = +/- 1 SE. *$p < 0.05$.

Behavioral categorization following training

We then assessed training-related improvements in categorization via listeners' identification of the musical interval continuum. We found a group x session interaction for identification slopes ($F_{1,26} = 4.93$, $p = 0.035$). Importantly, control and learning groups did not differ at pretest (**Fig. 3C**; $t_{26} = -0.14$, $p = 0.48$), suggesting common baseline categorization. Critically, post hoc analyses revealed that identification slopes were steeper at posttest for successful learners (**Fig. 3D-E**; $t_{26} = 4.42$, $p < 0.001$), whereas performance remain static in the control group ($t_{26} = 4.42$, $p = 0.21$). For learners, in addition to training gains (main effect of phase: $F_{1,13} = 11.65$, $p = 0.005$), achieving better accuracy during training was associated with steeper identification functions overall ($F_{1,13} = 8.58$, $p = 0.012$). Similarly, RTs showed a group x phase interaction ($F_{1,78} = 3.98$, $p = 0.050$). Whereas the control group achieved faster RTs at posttest ($t_{78} = -3.64$, $p < 0.001$), RTs remained constant in the learning group ($t_{78} = -0.73$, $p = 0.47$).

Electrophysiological results

ERP waveforms are shown per group and experimental phase in **Figure 4** (pooling all tokens). For the learning group, we found a training accuracy x phase interaction in P2 ($F_{1,39} = 5.77$, $p = 0.021$) and P1 amplitudes ($F_{1,39} = 11.29$, $p = 0.002$); better performance during training was associated with decreased amplitudes in the posttest but not the pretest (P2 posttest: $t_{39} = -3.41$, $p = 0.010$; P1 posttest: $t_{39} = -2.32$, $p = 0.010$). All other ERP comparisons with training accuracy were not significant.

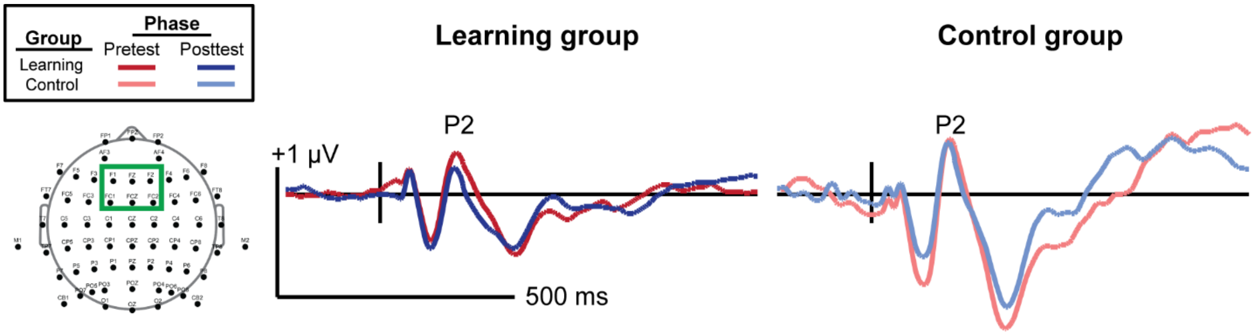


Figure 4: Grand average ERP waveforms collapsed across all tokens from the frontocentral electrode cluster (mean F1, Fz, F2, FC1, FCz, FC2; inset). The learning group (left) underwent brief identification training whereas the control group (right) did not. The tick mark represents t = 0 (stimulus onset).

In learners, we found an identification slopes x phase interaction for P2 amplitudes ($F_{1,38}$ = 4.16, p = 0.048); steeper (i.e., more categorical) posttest identification slopes were associated with a decrease in neural activity after training (**Fig. 5A**). Main effects of slope ($F_{1,39}$ = 8.46, p = 0.006) and phase ($F_{1,39}$ = 6.26, p=0.017) were also found for the slow wave (300-500 ms). Critically, these brain-behavior relationships were specific to learners and were not observed in the control group (**Fig. 5B**; all *p's* > 0.05).

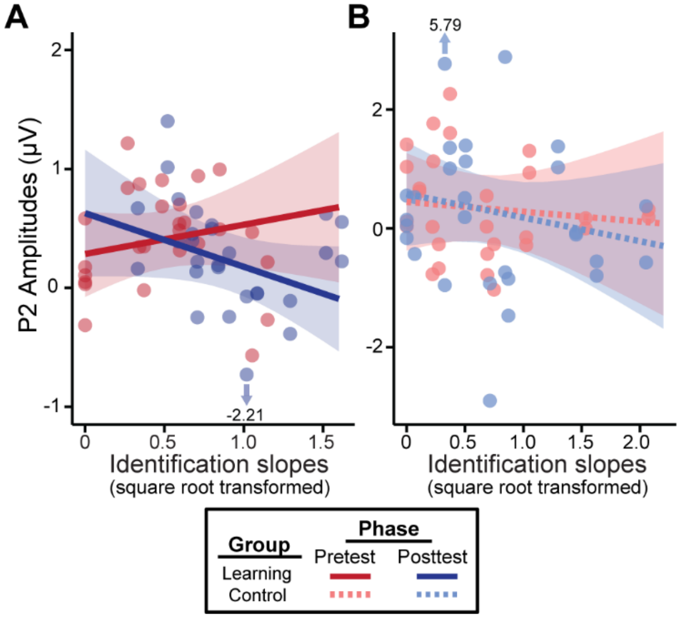


Figure 5: Neural amplitudes scale with behavioral outcomes in the learning group (A) but not the control group (B). Better posttest categorization (i.e., steeper identification slopes) is associated with a decrease in P2 amplitudes. Data points indicate individual subjects (collapsed across tokens 1 & 5 and 3). Arrows/values mark outliers (which did not alter results). Shading = 95% CI.

Hemispheric asymmetries were assessed via difference waveforms (i.e., mean[tokens 1 & 5] vs. 3) indexing the degree of categorization contained in neural responses. This analysis focused on electrodes T7 and T8 located over the left and right temporal lobes, respectively. We used a running paired t-test to evaluate training effects in a point-by-point manner across the ERP time courses (BESA Statistics, v2; **Fig. 6**). This revealed that in learners, category differentiation was modulated by learning 112-356 ms after stimulus onset. Training effects were most prominent over electrode T8 (right hemisphere; **Fig. 6B**). Guided by these results, we then extracted average amplitudes within this time window and ran a three-way mixed model ANOVA (group, identification slopes, training phase). The group x slope interaction was significant for electrode T8 ($F_{1,23} = 7.86$, $p = 0.010$). Post hoc analyses revealed that for learners, steeper identification slopes predicted larger (i.e., more categorical) responses over the right hemisphere ($t_{23} = 0.59$, $p = 0.021$). This brain-behavior relationship was not observed in controls nor over the left hemisphere (*p's* > 0.05). These data reveal a right hemisphere bias in neural mechanisms supporting category learning of musical sounds.

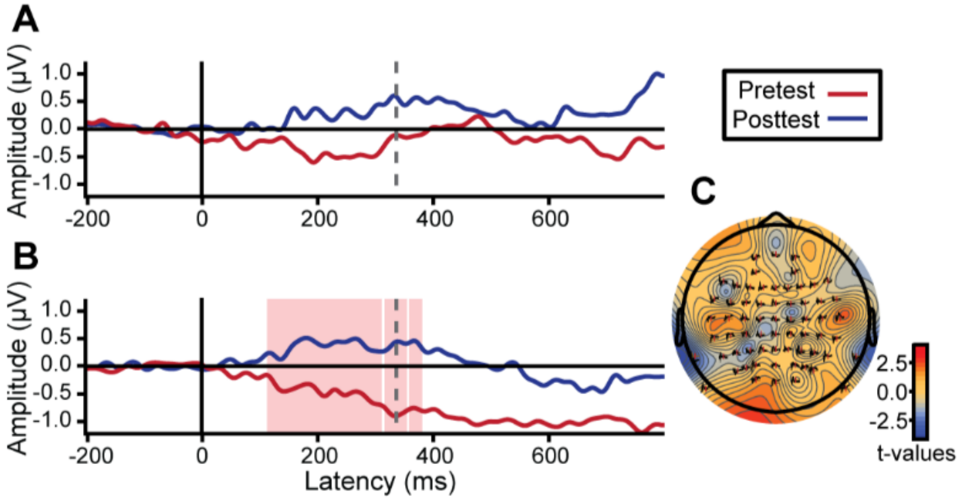


Figure 6: Neuroplastic changes following auditory categorical learning of music intervals are biased toward right hemisphere. Only data for the learning group is shown. (A-B) Difference waves (i.e., mean[token 1/5] – token 3) indexing categorical neural coding. An increase in neural categorization after training occurs over right (B; electrode T8) but not left hemisphere (A; electrode T7). Shaded region indicates a significant session effect ($p < 0.05$). (C) Topographic statistical map at t = 336 ms (dotted gray line in A & B) where pre- to post-test changes in categorical coding is maximal.

Neuroanatomical results

We next determined whether preexisting structural asymmetries (i.e., gray matter volume, cortical thickness) of primary auditory cortex were also associated with successful category learning. Volumetric analyses revealed that gray matter volumes were larger on average in the right compared to left PAC ($t_{11} = 12.36$, $p < 0.001$). The interaction of phase and structural measures were not significant for identification slopes. However, phase was kept in the models to isolate the relationship between structural PAC measures and behavior after factoring out training effects (see section *3.2*). Smaller gray matter volumes in right PAC were associated with stronger categorization overall ($F_{1,11} = 5.80$, $p = 0.035$, after accounting for effects of phase) (**Fig. 7**). Meanwhile, thinner cortical thickness of left PAC corresponded to better identification slopes ($F_{1,11} = 15.07$, $p = 0.003$, after accounting for effects of phase). Cortical thicknesses and gray matter volumes did not correlate with each other for either right or left PAC suggesting these volumetrics provided independent measures of the anatomy (all *p's* > 0.05). Taken together, these results indicate that preexisting differences in bilateral PAC structure predict individual categorization performance.

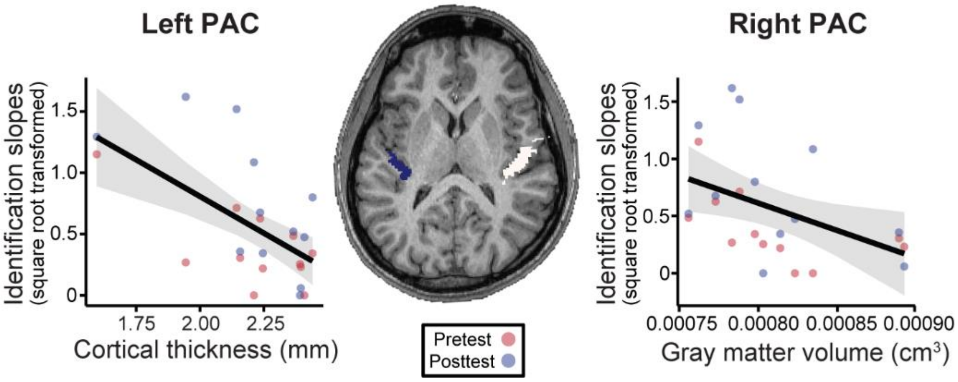


Figure 7: Neuroanatomical measures in primary auditory cortex (PAC) predict behavioral categorization performance in the learning group. (Left) In left HG, larger cortical thickness is associated with poorer categorization. (Right) Similarly, larger gray matter volumes (cm^3) in right HG were associated with poorer behavioral categorization. (Center) MRI image from a representative subject with left and right HG shown in blue and white, respectively. Data points indicate individual subject identification slopes. Shading = 95% CI.

Discussion

By measuring multichannel EEGs and brain volumetrics during short term auditory category learning tasks, our data reveal three primary findings: (i) rapid label learning of nonspeech sounds emerges very early in the brain (~150-200 ms, P2 wave), (ii) ERP responses decrease with more successful learning suggesting more efficient neural processing (i.e., reduced amplitudes) after training; (iii) neuroplastic changes in categorizing musical sounds are stronger in right hemisphere where smaller and thinner auditory cortical regions predicted better categorization performance. Successful category learning is therefore characterized by increased functional efficiency of sensory processing layered on preexisting structural advantages within auditory cortex.

Functional correlates of auditory category learning

Our data suggest category acquisition for nonspeech sounds is associated with changes in ERP P2. The functional significance of P2 is still poorly understood (Crowley & Colrain, 2004). Experience-dependent neuroplasticity in P2 has been interpreted as reflecting enhanced perceptual encoding and/or auditory object representations (Bidelman & Lee, 2015; Bidelman et al., 2014; Garcia-Larrea et al., 1992; Ross et al., 2013; Shahin et al., 2003), improvements in the task acquisition process (Tremblay et al., 2014), reallocation of attentional resources (Alain et al., 2007), increased inhibition of task-irrelevant signals (Seppanen et al., 2012; Sheehan et al., 2005), or mere stimulus exposure (Ross et al., 2013; Sheehan et al., 2005). Here, we demonstrate early ERP waves including P1 (~40-80 ms) as well as P2 (~150-200 ms) closely scale with behavioral learning. Moreover, these neuroplastic effects are surprisingly fast, occurring rapidly within only 20 minutes of training. Our findings parallel visual category learning where changes in the visual-evoked N1 and late positive component signal successful learning (Perez-Gay

Juarez et al., 2019). Our results also align with previous studies using various auditory training tasks including speech (Alain et al., 2010; Alain et al., 2007; Ben-David et al., 2011; Tremblay et al., 2001; Tremblay & Kraus, 2002; Tremblay et al., 2009) and nonspeech sounds (Atienza et al., 2002; Bosnyak et al., 2004; Tong et al., 2009; Wisniewski et al., 2020) suggesting P2 indexes auditory experience that reflects learning success and is not simply a product of the task acquisition process (cf. Tremblay et al., 2014) or repeated stimulus exposure (Ross et al., 2013; Ross & Tremblay, 2009; Sheehan et al., 2005). The lack of clear neural changes in control listeners also contradicts exposure or repetition effect interpretations of our data.

In this study, successful learning (i.e., both training accuracy and identification function slopes) was characterized by a *reduction* in ERP amplitudes after training. The specific direction of P2 modulations varies across experiments with some reporting an increase in evoked responses with learning (Atienza et al., 2002; Bosnyak et al., 2004; Carcagno & Plack, 2011; Ross et al., 2013; Sheehan et al., 2005; Tong et al., 2009; Tremblay et al., 2001; Wisniewski et al., 2020) and others a decrease (Alain et al., 2010; Ben-David et al., 2011; Zhang et al., 2005). As suggested by Alain et al. (2010), such discrepancies could be related to the task (e.g., active task vs. passive recording), the stimuli (e.g., speech vs. nonspeech), the rate of learning among the participants, or even the rigor of training paradigm. Studies reporting enhanced P2 often included multiple days of training or recorded ERPs during passive listening (Atienza et al., 2002; Bosnyak et al., 2004; Ross et al., 2013; Seppanen et al., 2013; Tremblay et al., 2001; Wisniewski et al., 2020). Long-term auditory experiences (e.g., music training, tone language expertise) have also been associated with enhanced P2 during active sound categorization (Bidelman & Alain, 2015; Bidelman & Lee, 2015; Bidelman et al., 2014) as well as learning (Seppanen et al., 2012, 2013; Shahin et al., 2003). The ERP decreases we find in successful learners are highly consistent with single-session, rapid learning experiments demonstrating greater efficiency of sensory-evoked neural responses during active task engagement (Alain et al., 2010; Ben-David et al., 2011; Guenther et al., 2004; Perez-Gay Juarez et al., 2019; Sohoglu & Davis, 2016). Consequently, our results reinforce notions that the P2 is a biomarker of learning to classify auditory stimuli and map sounds to labels (Ben-David et al., 2011; Rosen & Howell, 1987; cf. Tremblay et al., 2014).

On the contrary, decreased neural activity might reflect other aspects of the task, including arousal and/or fatigue (Crowley & Colrain, 2004; Näätänen & Picton, 1987). However, decreased neural activity from these factors would have been expected in both groups due to the similar task constraints on all participants. If better learners simply sustain arousal more effectively through posttest, we would have also expected faster RTs. Rather, our data suggest decreases in activation meaningfully reflect music category learning, paralleling findings with speech (Guenther et al., 2004). Alternatively, given modulations in both P2 and slow wave activity, a separate but overlapping processing negativity within this timeframe cannot be ruled out. Negative processing components have been associated with early auditory selection and attention (Crowley & Colrain, 2004; Hillyard & Kutas, 1983; Näätänen & Picton, 1987) and may therefore be another target for learning processes.

Hemispheric lateralization and music categorization

Our findings show that acquiring novel categories for musical intervals predominantly recruits neural resources from the right auditory cortex, complementing the left hemisphere bias reported for speech categorization (Alho et al., 2016; Chang et al., 2010; Desai et al., 2008; Golestani & Zatorre, 2004; Liebenthal et al., 2005; Liebenthal et al., 2010; Myers et al., 2009; Zatorre et al., 1992). Specifically, we observed enhanced neural categorization over the right hemisphere in more successful learners. Gray matter volumetrics in right PAC were also associated with behavioral categorization abilities. These findings support long-standing notions about lateralization for speech vs. music categorization in the brain (Alho et al., 2016; Bidelman & Walker, 2019; Bouton et al., 2018; Chang et al., 2010; Desai et al., 2008; Klein & Zatorre, 2011, 2015; Liebenthal et al., 2010; Mankel, Barber, et al., 2020; Zatorre et al., 1992). Superior music categorization in both trained musicians (Bidelman & Walker, 2019; Klein & Zatorre, 2011, 2015) as well as musically adept non-musicians (Mankel, Barber, et al., 2020) has been associated with right temporal lobe functions. We thus provide new evidence that even brief, 20-minute identification training is sufficient to recruit right hemisphere neural circuity that subserves auditory sensory coding and classification of musical stimuli.

Neuroanatomical correlates of auditory category learning

Our MRI results demonstrate that individual variation in structural measures (gray matter volume, cortical thickness) also predict behavioral categorization performance beyond mere training effects. Brain structure is influenced by genetic, epigenetic, and experiential factors (Zatorre et al., 2012). Thus, it is often difficult to know whether anatomical differences are innate or experience-driven, but structural measures are presumed to be more stable and less plastic than functional responses (e.g., ERPs) (Golestani, 2012). Structural plasticity associated with music training, for example, may take weeks or months to appear (Wenger et al., 2020), suggesting the results observed in this study reflect preexisting neuroanatomical differences rather than structural changes due to the brief identification training. Anatomical variability in auditory cortex has been related to learning rate and attainment for foreign speech sounds (Golestani et al., 2007), linguistic pitch patterns (Wong et al., 2008), and melody discrimination (Foster & Zatorre, 2010a) as well as native speech categorization (Fuhrmeister & Myers, 2021).

Consistent with this prior work on speech, our findings suggest that individual differences in music category perception and functional plasticity are influenced by anatomical predispositions within auditory cortex—that is, a layering of both nature and nurture.

It is often assumed larger morphology within a particular brain area yields better computational efficiency (i.e., “bigger is better”) (Kanai & Rees, 2011). For example, faster, more successful learners of nonnative speech sounds show more voluminous primary auditory cortex and adjacent white matter in left hemisphere (Golestani et al., 2007; Golestani et al., 2002; Wong et al., 2008). Relatedly, expert listeners (i.e., musicians) often show increased gray matter volumes and cortical thickness in PAC (Bermudez et al., 2009; de Manzano & Ullen, 2018; Gaser & Schlaug, 2003; Schneider et al., 2002; Seither-Preisler et al., 2014; Wengenroth et al., 2014; cf. Wenger et al., 2020). Instead, our data show the opposite pattern with regard to nonspeech category learning. To our knowledge, only one study has shown correspondence between decreased gyrification in temporal regions and improved consistency in speech categorization behaviors (Fuhrmeister & Myers, 2021). Similarly, smaller gray matter volume in STG has been linked to improvements in speech and cognitive training (Maruyama et al., 2018; Takeuchi, Taki, Hashizume, et al., 2011; Takeuchi, Taki, Sassa, et al., 2011). Congenital amusia—a disorder characterized by impairments in the perception and/or production of music—has also been associated with increased gray matter volume and cortical thickness along the frontotemporal neural pathway (Hyde et al., 2007; Hyde et al., 2006). Thus, it seems “less is more” with respect to the expanse of auditory anatomy and certain aspects of listening performance. However, future research is needed to clarify the relationships between macroscopic gray and white matter volumes measured by MRI, neuronal microstructures, and their behavioral correlates.

Conclusion

We demonstrate that rapid auditory category learning of musical interval sounds is associated with increased efficiency in sensory processing in bilateral, though predominantly right, auditory cortex. The relationship between better behavioral gains in identification performance and the ERPs corroborate P2 as an index of auditory experience and a biomarker for successful perceptual learning. The right hemisphere dominance supporting music category learning complements left hemisphere networks reported for speech categorization. Individual categorization performance is therefore characterized by short-term functional changes superimposed on preexisting structural differences in bilateral auditory areas.

Chapter 4

General Discussion

The results of these studies demonstrate that auditory categorization is a complex interaction between preexisting and experience-dependent functional differences in the brain, and their relative influence on categorical functions predictably varies across individuals. The first study of this dissertation showed that inherent musicality—musical listening abilities distinct from basic psychoacoustic measures or those gained during formal music training—was associated with more efficient categorical processing (i.e., faster reaction times) and more robust neural differentiation of speech categories. The second study revealed the acquisition of novel, nonspeech auditory categories is a layering of neuroplastic changes in sensory function coupled with structural advantages for auditory processing. We thus provide evidence that categorical perception is neither solely pre-determined at birth nor exclusively developed through learning and experience but rather a highly individualized product of both nature and nurture factors.

Both studies demonstrated that certain individuals exhibit natural sensory processing advantages within the auditory system which enables more efficient and robust categorization. The "musical sleepers" in Chapter 2 (i.e., nonmusicians with inherently superior musical listening abilities; Law & Zentner, 2012) were faster overall in identifying vowel categories. These individuals also showed more categorical event-related potentials (ERPs) compared to listeners with poorer musical-listening skills, specifically better P2-latency differentiation between prototypical and ambiguous vowels. As shown in Chapter 3, successful learners of musical interval categories (i.e., training accuracy and posttest identification function slopes) displayed more efficient early sensory ERP responses (i.e., P1, P2) after identification training. Although the functional significance of P2 is still poorly understood (Crowley & Colrain, 2004), prior work has suggested that P2 reflects pre-cognitive (and pre-motor) perceptual encoding, formation of auditory object representations, and stimulus classification or evaluation of stimulus ambiguity within the brain (Bidelman & Lee, 2015; Bidelman et al., 2014; Garcia-Larrea et al., 1992; Ross et al., 2013; Shahin et al., 2003; Silva et al., 2020). The results of this dissertation are consistent with these interpretations and extend them by demonstrating that (i) P2 is sensitive to individual differences in categorical perception; (ii) P2 indexes successful learning performance during the acquisition of auditory categories; and (iii) superior categorization abilities are characterized by more efficient pre-attentive, sensory-encoding processes within auditory cortex.

Another common finding across these studies was the prevalence of laterality effects—mainly right hemisphere—in both better speech categorizers and music interval category learners. Categorical processing of speech and music are thought to primarily invoke left versus right hemisphere dominance in the brain, respectively (Alho et al., 2016; Bidelman & Walker, 2019; Bouton et al., 2018; Chang et al., 2010; Desai et al., 2008; Klein & Zatorre, 2011, 2015; Liebenthal et al., 2010; Mankel, Barber, et al., 2020; Zatorre et al., 1992). Yet, listeners with highly adept (musical) listening abilities showed more robust categorical processing of speech in the right auditory cortex. As proposed in Chapter 2, these musical sleepers may recruit additional resources within the right hemisphere for speech perception which could contribute to more efficient ERP latencies and reaction times. We also observed enhanced neural categorization over the right hemisphere in more successful category learners, and gray matter volumes of right auditory cortex predicted better categorization abilities in Chapter 3. An alternative possibility is that categorical processing is relatively stable across individuals in the left hemisphere due to either less morphological variability than the right auditory cortex (Penhune et al., 1996; Westbury et al., 1999) or the overlearned nature of native speech (Bulgarelli & Weiss, 2016), thus permitting greater sensitivity to individual differences in categorization abilities within the right hemisphere.

Finally, the broader theme of this dissertation concerned the relative contribution of nature versus nurture factors in auditory categorization. Chapter 2 examined these factors through the lens of musicality. Although musicality is a multifaceted, complex phenotype that is influenced by both genetics and environmental influences, (Gingras et al., 2015; Müllensiefen et al., 2014; Swaminathan & Schellenberg, 2018; Ullén et al., 2014), it is often presumed that performance on musical listening tests reflects “natural” or “innate” differences in musical ability (Mankel & Bidelman, 2018; Swaminathan & Schellenberg, 2018). While musicians show enhanced neural processing of speech and music (e.g., Bidelman & Alain, 2015; Bidelman & Krishnan, 2010; Bidelman & Walker, 2019; Bidelman et al., 2014; Kraus et al., 2009; Marie et al., 2011; Nan et al., 2018; Parbery-Clark et al., 2009; Shahin et al., 2003), formal music training itself is not a necessary (though perhaps sufficient) condition to enhance the categorization of complex sounds (Mankel & Bidelman, 2018). An unanswered question stemming from this research is whether the presumed music training-induced neuroplasticity and widely reported “musician advantage” in auditory tasks is actually due to intrinsic perceptual processing differences that motivate long-term engagement with formal music experiences.

The results in Chapter 3 further suggest that individual differences in categorical perception and functional plasticity are influenced by inherent structural differences in auditory cortex. Specifically, thinner and smaller auditory cortical regions predicted better categorization performance. Like musicality traits, brain structure is shaped by genetic, epigenetic, and experiential factors (Zatorre et al., 2012). However, structural measures are presumed to be more stable and less plastic than functional (ERP) responses (Golestani, 2012), and any individual differences in brain structure observed here are likely innate or emerge very early in the neurodevelopmental timeline (see Fuhrmeister & Myers, 2021) rather than from the training itself. We therefore concluded that successful category learning is driven by a layering of functional, neuroplastic changes superimposed on preexisting structural differences within auditory cortex.

Limitations

We acknowledge that it is difficult to tease apart truly innate predispositions from experience-dependent factors in auditory system function and structure. It is possible that the individuals across these studies differed on some other environmental or experiential factor rather than inherent auditory-perceptual abilities, per se. We cannot rule out additional influences such as early exposure to music (Tierney et al., 2015), daily recreational, or personal engagement with music listening (Müllensiefen et al., 2014) or other auditory stimuli that possibly contributed to higher PROMS scores and auditory processing advantages in otherwise musically naïve listeners (Mankel & Bidelman, 2018). However, animal studies show mere passive exposure to auditory stimuli is insufficient to induce neuroplastic changes in primary auditory cortex (Fritz et al., 2005). It is therefore unlikely that prior informal music exposure alone drives the enhancements we observed in high performing listeners.

Given that the learners in the second study had additional training trials, it could also be argued that the observed neural changes in this group are merely due to increased exposure to the musical interval stimuli relative to controls (Ross & Tremblay, 2009; Sheehan et al., 2005; Tremblay et al., 2010). Yet, the relationship between individual categorization performance and our neural measures suggests these differences were behaviorally meaningful during category identification and/or the acquisition of novel categories. Our results are consistent with the notion that neural correlates of categorical perception emerge when listeners are actively engaged with in categorizing familiar (or learned) auditory categories (i.e., not during passive listening; Bidelman & Walker, 2017). Additionally, ERP responses scale with individual performance in the learning group—where the number of trials was similar—further refuting a mere exposure-based explanation.

Conclusion

Collectively, this dissertation work proposes that auditory categorization is a complex interaction of both nature and nurture effects. Inherent differences in neural function foster sensory processing advantages within auditory cortex that are shaped by experience (e.g., music training; learning) and promote enhanced categorization and perception of sounds. Our results suggest that the auditory cortex is a source of neural variability for both categorical speech perception and nonspeech category learning.

Future directions

These findings generate new avenues for uncovering the fundamental neural mechanisms of category acquisition and auditory learning. Planned future studies will incorporate source analysis (see Chapter 2) and connectivity techniques to determine whether auditory category learning is a bottom-up, sensory-driven or top-down cognitive process. Extensions of this work should also evaluate whether the neural targets identified for successful nonspeech category learning predict learning in other domains, such as nonnative phoneme identification (Kajiura et al., 2021; Wong et al., 2007) or reading abilities (e.g., Chiappe et al., 2001). Our research also highlights the need to assess preexisting behavioral and neural factors before attributing enhanced auditory processing effects to neuroplastic changes engendered by other learning activities such as music or cognitive training (Foroughi et al., 2016; Mankel & Bidelman, 2018).

Additionally, further research is necessary to determine whether deficits in speech categorization, such as those observed in dyslexia or specific language impairment (Brandt & Rosen, 1980; Coady et al., 2005; Joanisse et al., 2000; Noordenbos & Serniclaes, 2015; Werker & Tees, 1987), are merely part of a normal distribution (albeit more extreme values) of individual differences in sensory, phonological, or even lexical processes (McMurray et al., 2010). Alternatively, future efforts may uncover fundamental neurophysiological impairments that distinguish neurotypical, normal-hearing individual differences and those with speech-language disorders. A better understanding of the neurobiology supporting categorization abilities across normal and disordered populations could potentially lead to more tailored rehabilitation and training programs that better maximize therapeutic or learning benefits for reading, language, and/or receptive hearing skills.

References

Al-Fahad, R., Yeasin, M., & Bidelman, G. M. (2020). Decoding of single-trial EEG reveals unique states of functional brain connectivity that drive rapid speech categorization decisions. *Journal of Neural Engineering, 17*(1), 016045. https://doi.org/10.1088/1741-2552/ab6040

Alain, C., Campeanu, S., & Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *Journal of Cognitive Neuroscience, 22*(2), 392-403. https://doi.org/10.1162/jocn.2009.21279

Alain, C., Snyder, J. S., He, Y., & Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cerebral Cortex, 17*(5), 1074-1084. https://doi.org/10.1093/cercor/bhl018

Alho, J., Green, B. M., May, P. J. C., Sams, M., Tiitinen, H., Rauschecker, J. P., & Jääskeläinen, I. P. (2016). Early-latency categorical speech sound representations in the left inferior frontal gyrus. *Neuroimage, 129*, 214-223. https://doi.org/10.1016/j.neuroimage.2016.01.01

Atienza, M., Cantero, J. L., & Dominguez-Marin, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning & Memory, 9*(3), 138-150. https://doi.org/10.1101/lm.46502

Başkent, D., & Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *Journal of the Acoustical Society of America, 139*(3), EL51-EL56. https://doi.org/doi:http://dx.doi.org/10.1121/1.4942628

Ben-David, B. M., Campeanu, S., Tremblay, K. L., & Alain, C. (2011). Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. *Psychophysiology, 48*(6), 797-807. https://doi.org/10.1111/j.1469-8986.2010.01139.x

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. https://doi.org/10.1093/cercor/bhn196

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology, 2*, 94. https://doi.org/10.3389/fpsyg.2011.00094

Bidelman, G. M. (2015). Induced neural beta oscillations predict categorical speech perception abilities. *Brain and Language, 141*, 62-69. https://doi.org/http://dx.doi.org/10.1016/j.bandl.2014.11.003

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *Journal of Neuroscience, 35*(2), 1240 –1249. https://doi.org/10.1523/jneurosci.3292-14.2015

Bidelman, G. M., Bush, L. C., & Boudreaux, A. M. (2020). Effects of Noise on the Behavioral and Neural Categorization of Speech. *Frontiers in Neuroscience, 14*, 153. https://doi.org/10.3389/fnins.2020.00153

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS One, 8*(4), e60676. https://doi.org/10.1371/journal.pone.0060676

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research, 1355*, 112-125. https://doi.org/10.1016/j.brainres.2010.07.100

Bidelman, G. M., & Lee, C.-C. (2015). Effects of language experience and stimulus context on the neural organization and categorical perception of speech. *Neuroimage, 120*, 191-200. https://doi.org/10.1016/j.neuroimage.2015.06.087

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage, 79*(1), 201-212. https://doi.org/10.1016/j.neuroimage.2013.04.093

Bidelman, G. M., Sigley, L., & Lewis, G. (2019). Acoustic noise and vision differentially warp speech categorization. *Journal of the Acoustical Society of America*. https://doi.org/10.1121/1121.5114822

Bidelman, G. M., & Walker, B. S. (2017). Attentional modulation and domain-specificity underlying the neural organization of auditory categorical perception. *European Journal of Neuroscience, 45*(5), 690-699. https://doi.org/10.1111/ejn.13526

Bidelman, G. M., & Walker, B. S. (2019). Plasticity in auditory categorization is supported by differential engagement of the auditory-linguistic network. *Neuroimage, 201*, 1-10. https://doi.org/10.1101/663799

Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience, 40*, 2662-2673. https://doi.org/10.1111/ejn.12627

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience, 7*(3), 295-301. https://doi.org/10.1038/nn1198

Bosnyak, D. J., Eaton, R. A., & Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex, 14*(10), 1088-1099. https://doi.org/10.1093/cercor/bhh068

Bouton, S., Chambon, V., Tyrand, R., Guggisberg, A. G., Seeck, M., Karkar, S., van de Ville, D., & Giraud, A. L. (2018). Focal versus distributed temporal cortex activity for speech sound category assignment. *Proceedings of the National Academy of Sciences of the United States of America, 115*(6), E1299-E1308. https://doi.org/10.1073/pnas.1714279115

Brandt, J., & Rosen, J. J. (1980). Auditory phonemic perception in dyslexia: Categorical identification and discrimination of stop consonants. *Brain and Language, 9*(2), 324-337.

Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of overlearning on entrenchment effects in statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1621-1631. https://doi.org/10.1037/xlm0000263

Burns, E. M., & Ward, W. D. (1978). Categorical perception - phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America, 63*(2), 456-468.

Carcagno, S., & Plack, C. J. (2011). Pitch discrimination learning: Specificity for pitch and harmonic resolvability, and electrophysiological correlates. *Journal of the Association for Research in Otolaryngology, 12*, 503-517. https://doi.org/10.1007/s10162-011-0266-3

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience, 13*(11), 1428-1432. https://doi.org/10.1038/nn.2641

Chiappe, P., Chiappe, D. L., & Siegel, L. S. (2001). Speech perception, lexicality, and reading skill. *Journal of Experimental Child Psychology, 80*(1), 58-74. https://doi.org/10.1006/jecp.2000.2624

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108*(3), 804-809. https://doi.org/10.1016/j.cognition.2008.04.004

Coady, J. A., Kluender, K. R., & Evans, J. L. (2005). Categorical perception of speech by children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 48*(4), 944-959. https://doi.org/10.1044/1092-4388(2005/065)

Crowder, R. G. (1989). Categorical perception of speech: A largely dead horse, surpassingly well kicked. *Behavioral and Brain Sciences, 12*(4), 760.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115*, 732–744. https://doi.org/10.1016/j.clinph.2003.11.021

Damper, R. I., & Harnad, S. R. (2000). Neural network models of categorical perception. *Perception and Psychophysics, 62*(4), 843-867.

Das, S. R., Avants, B. B., Grossman, M., & Gee, J. C. (2009). Registration based cortical thickness measurement. *Neuroimage, 45*(3), 867-879. https://doi.org/10.1016/j.neuroimage.2008.12.016

de Manzano, O., & Ullen, F. (2018). Same Genes, Different Brains: Neuroanatomical Differences Between Monozygotic Twins Discordant for Musical Training. *Cerebral Cortex*, *28*(1), 387-394. https://doi.org/10.1093/cercor/bhx299

de Souza, A. C. S., Yehia, H. C., Sato, M.-a., & Callan, D. (2013). Brain activity underlying auditory perceptual learning during short period training: simultaneous fMRI and EEG recording. *BMC Neuroscience*, *14*, 8. https://doi.org/10.1186/1471-2202-14-8

Desai, R., Liebenthal, E., Waldron, E., & Binder, J. R. (2008). Left posterior temporal regions are sensitive to auditory categorization. *Journal of Cognitive Neuroscience*, *20*(7), 1174-1188. https://doi.org/10.1162/jocn.2008.20081

Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0805022105

Dittinger, E., D'Imperio, M., & Besson, M. (2018). Enhanced neural and behavioural processing of a nonnative phonemic contrast in professional musicians. *European Journal of Neuroscience, 47*(12), 1504-1516. https://doi.org/10.1111/ejn.13939

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Elmer, S., Meyer, M., & Jancke, L. (2012). Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cerebral Cortex*, *22*(3), 650-658. https://doi.org/10.1093/cercor/bhr142

Entwislea, D. R., & Astone, N. M. (1994). Some Practical Guidelines for Measuring Youth's Race/Ethnicity and Socioeconomic Status. *Child Development, 65*(6). https://doi.org/10.1111/j.1467-8624.1994.tb00833.x

Escudero, P., Benders, T., & Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second language vowels. *Journal of the Acoustical Society of America*, *130*(4), EL206-212. https://doi.org/10.1121/1.3629144

Evans, A. C., Collins, D. L., Mills, S., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. 1993 IEEE conference record nuclear science symposium and medical imaging conference,

Feng, G., Gan, Z., Wang, S., Wong, P. C. M., & Chandrasekaran, B. (2017). Task-General and Acoustic-Invariant Neural Representation of Speech Categories in the Human Brain. *Cerebral Cortex*, *28*(9), 3241–3254. https://doi.org/10.1093/cercor/bhx195

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*, 1-34. https://doi.org/10.1017/S0272263104261010

Foroughi, C. K., Monfort, S. S., Paczynski, M., McKnight, P. E., & Greenwood, P. M. (2016). Placebo effects in cognitive training. *Proceedings of the National Academy of Sciences of the United States of America, 113*(27), 7470-7474. https://doi.org/10.1073/pnas.1601243113

Foster, N. E., & Zatorre, R. J. (2010a). Cortical structure predicts success in performing musical transformation judgments. *Neuroimage, 53*(1), 26-36. https://doi.org/10.1016/j.neuroimage.2010.06.042

Foster, N. E., & Zatorre, R. J. (2010b). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex, 20*(6), 1350-1359. https://doi.org/10.1093/cercor/bhp199

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36*, 268-294. https://doi.org/10.1016/j.wocn.2007.06.005

Fritz, J., Elhilali, M., & Shamma, S. (2005). Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex [Article; Proceedings Paper]. *Hearing Research, 206*(1-2), 159-176. https://doi.org/10.1016/j.heares.2005.01.015

Fuhrmeister, P., & Myers, E. B. (2021). Structural neural correlates of individual differences in categorical perception. *Brain and Language, 215*. https://doi.org/10.1016/j.bandl.2021.104919

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110-125.

Garcia-Larrea, L., Lukaszewicz, A. C., & Mauguiere, F. (1992). Revisiting the oddball paradigm: Non-target vs. neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia, 30*(8), 723-741. https://doi.org/10.1016/0028-3932(92)90042-k

Gaser, C., & Schlaug, G. (2003). Gray matter differences between musicians and nonmusicians. *Annals of New York Academy of Science, 999*, 514-517. https://doi.org/10.1196/annals.1284.062

Gerrits, E., & Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Perception and Psychophysics, 66*(3), 363-376. https://doi.org/10.3758/BF03194885

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society B-Biological Sciences, 370*(1664), 76-90. https://doi.org/10.1098/rstb.2014.0092

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(1), 69-78. https://doi.org/10.1002/wcs.26

Goldstone, R. L., Steyvers, M., & Larimer, K. (1996). Categorical perception of novel dimensions. Eighteenth Annual Conference of the Cognitie Science Society, Hillsdale, NJ.

Golestani, N. (2012). Brain structural correlates of individual differences at low-to high-levels of the language processing hierarchy: A review of new approaches to imaging research. *International Journal of Bilingualism*, *18*(1), 6-34. https://doi.org/10.1177/1367006912456585

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, *17*(3), 575-582. https://doi.org/10.1093/cercor/bhk001

Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron*, *35*, 997-1010. https://doi.org/10.1016/S0896-6273(02)00862-0

Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: Reallocation of neural substrates. *Neuroimage*, *21*(2), 494-506. https://doi.org/10.1016/j.neuroimaging.2003.09.071

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*. https://doi.org/10.1016/j.bandl.2008.01.005

Gordon, E. E. (1965). *Musical aptitude profile manual*. Houghton Mifflin.

Goudbeek, M., Swingley, D., & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1913-1933. https://doi.org/10.1037/a0015781

Guenther, F. H., Husain, F. T., Cohen, M. A., & Shinn-Cunningham, B. G. (1999). Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America*, *106*(5), 2900-2912. https://doi.org/10.1121/1.428112

Guenther, F. H., Nieto-Castanon, A., Ghosh, S. S., & Tourville, J. A. (2004). Representation of sound categories in auditory cortical maps. *Journal of Speech, Language, and Hearing Research*, *47*(1), 46-57. https://doi.org/10.1044/1092-4388(2004/005)

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107*(5 Pt 1), 2711-2724. https://doi.org/10.1121/1.428657

Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2016). Music and dyslexia: A new musical training method to improve reading and related disorders. *Frontiers in Psychology*, *7*, 26. https://doi.org/10.3389/fpsyg.2016.00026

Harnad, S. (1987). *Categorical perception: The groundwork of cognition*. Cambridge University Press.

Harnad, S., Hanson, S. J., & Lubin, J. (1991). Categorical perception and the evolution of supervised learning in neural nets. AAAI Spring Symposium on Machine Learning of Natural Language and Ontology, Stanford, CA.

Hazan, V., & Rosen, S. (1991). Individual variability in the perception of cues to place contrasts in initial stops. *Perception & Psychophysics, 49*(2), 187-200. https://doi.org/10.3758/bf03205038

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*, 33-61.

Howard, D., Rosen, S., & Broad, V. (1992). Major/Minor triad identification and discrimination by musically trained and untrained listeners. *Music Perception, 10*(2), 205–220. https://doi.org/10.2307/40285607

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *Journal of Neuroscience, 27*(47), 13028-13032. https://doi.org/10.1523/JNEUROSCI.3039-07.2007

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. *Brain, 129*(10), 2562-2570. https://doi.org/10.1093/brain/awl204

Idemaru, K., Holt, L. L., & Seltman, H. (2012). Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *Journal of the Acoustical Society of America, 132*(6), 3950-3964. https://doi.org/10.1121/1.4765076

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: speech perception, phonology, and morphology. *Journal of Experimental Child Psychology, 77*(1), 30-60. https://doi.org/10.1006/jecp.1999.2553

Joanisse, M. F., Zevin, J. D., & McCandliss, B. D. (2007). Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex, 17*(9), 2084-2093. https://doi.org/10.1093/cercor/bhl124

Kajiura, M., Jeong, H., Kawata, N. Y. S., Yu, S., Kinoshita, T., Kawashima, R., & Sugiura, M. (2021). Brain activity predicts future learning success in intensive second language listening training. *Brain and Language, 212*, 104839. https://doi.org/10.1016/j.bandl.2020.104839

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience, 12*(4), 231-242. https://doi.org/10.1038/nrn3000

Kapnoula, E., Edwards, J., & McMurray, B. (in press). Gradient activation of speech categories facilitates listeners' recovery from lexical garden paths, but not perception of speech-in-noise. *Journal of Experimental Psychology: Human Perception and Performance*.

Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J., & McMurray, B. (2017). Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance, 43*(9), 1594-1611. https://doi.org/10.1037/xhp0000410

Klatt, D. H. (1979). Speech perception: a model of acoustic–phonetic analysis and lexical access. *Journal of Phonetics*, *7*(3), 279-312. https://doi.org/10.1016/s0095-4470(19)31059-9

Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia*, *49*(5), 878-887. https://doi.org/10.1016/j.neuropsychologia.2011.01.008

Klein, M. E., & Zatorre, R. J. (2015). Representations of invariant musical categories are decodable by pattern analysis of locally distributed BOLD responses in superior temporal and intraparietal sulci. *Cerebral Cortex*, *25*(7), 1947-1957. https://doi.org/10.1093/cercor/bhu003

Kong, E. J., & Edwards, J. (2011). Individual differences in speech perception: Evidence from visual analogue scaling and eye-tracking. XVIIth International Congress of Phonetic Sciences, Hong Kong, China.

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, *59*, 40-57. https://doi.org/10.1016/j.wocn.2016.08.006

Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Annals of the New York Academy of Sciences*, *1169*, 543-557. https://doi.org/10.1111/j.1749-6632.2009.04549.x

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22-44.

Kruschke, J. K. (1993). Human Category Learning: Implications for Backpropagation Models. *Connection Science*, *5*(1), 3-36. https://doi.org/10.1080/09540099308915683

Kuhl, P., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), 13-21. https://doi.org/10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*(2), 93-107. https://doi.org/10.3758/BF03212211

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831-843. https://doi.org/10.1038/nrn1533

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 363*(1493), 979-1000. https://doi.org/10.1098/rstb.2007.2154

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science, 190*(4209), 69-72.

Kunert, R., Willems, R. M., & Hagoort, P. (2016). An Independent Psychometric Evaluation of the PROMS Measure of Music Perception Skills. *PLoS One, 11*(7), e0159103. https://doi.org/10.1371/journal.pone.0159103

Lane, H. (1965). The motor theory of speech perception: A critical review. *Psychological Review, 72*, 275-309.

Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: construction and validation of the Profile of Music Perception Skills. *PLoS One, 7*(12), 1-15. https://doi.org/10.1371/journal.pone.0052508.g001

Lee, R. R. W., Hsu, C. H., Lin, S. K., Wu, D. H., & Tzeng, O. J. L. (2017). Learning transforms functional organization for Mandarin lexical tone discrimination in the brain: Evidence from a MEG experiment on second language learning. *Journal of Neurolinguistics, 42*, 124-139. https://doi.org/10.1016/j.jneuroling.2016.12.004

Lee, Y.-S., Turkeltaub, P., Granger, R., & Raizada, R. D. S. (2012). Categorical speech processing in Broca's area: An fMRI study using multivariate pattern-based analysis. *Journal of Neuroscience, 32*(11), 3942-3948. https://doi.org/10.1523/jneurosci.3814-11.2012

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods, 38*(2), 202-210. https://doi.org/10.3758/BF03192770

Li, X. T., De Beuckelaer, A., Guo, J. H., Ma, F. L., Xu, M., & Liu, J. (2014). The Gray Matter Volume of the Amygdala Is Correlated with the Perception of Melodic Intervals: A Voxel-Based Morphometry Study. *PLoS One, 9*(6), Article e99889. https://doi.org/10.1371/journal.pone.0099889

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431-461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phonemic boundaries. *Journal of Experimental Psychology, 54*, 358-368.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243*, 489–494. https://doi.org/10.1126/science.2643163

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex, 15*(10), 1621-1631. https://doi.org/10.1093/cercor/bhi040

Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., & Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. *Cerebral Cortex*, *20*(12), 2958-2970. https://doi.org/10.1093/cercor/bhq045

Little, D. F., Cheng, H. H., & Wright, B. A. (2019). Inducing musical-interval learning by combining task practice with periods of stimulus exposure alone. *Attention, Perception, & Psychophysics 81*(1), 344-357. https://doi.org/10.3758/s13414-018-1584-x

Liu, R., & Holt, L. L. (2011). Neural changes associated with nonspeech auditory category learning parallel those of speech category acquisition. *Journal of Cognitive Neuroscience, 23*(3), 683-698. https://doi.org/10.1162/jocn.2009.21392

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242–1255.

Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology-Learning Memory and Cognition*, *24*(3), 732-753. https://doi.org/10.1037/0278-7393.24.3.732

Locke, S., & Kellar, L. (1973). Categorical perception in a non-linguistic mode. *Cortex, 9*(4), 355-369.

Luthra, S., Fuhrmeister, P., Molfese, P. J., Guediche, S., Blumstein, S. E., & Myers, E. B. (2019). Brain-behavior relationships in incidental learning of non-native phonetic categories. *Brain and Language, 198*, 104692. https://doi.org/10.1016/j.bandl.2019.104692

Mankel, K., Barber, J., & Bidelman, G. M. (2020). Auditory categorical processing for speech is modulated by inherent musical listening skills. *NeuroReport, 31*(2), 162-166. https://doi.org/10.1097/WNR.0000000000001369

Mankel, K., & Bidelman, G. M. (2018). Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proceedings of the National Academy of Sciences of the United States of America, 115*(51), 13129-13134. https://doi.org/10.1073/pnas.1811793115

Mankel, K., Pavlik Jr, P. I., & Bidelman, G. M. (2020). Single-trial neural dynamics influence auditory category learning. *bioRxiv [preprint]*.

Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2009). Open Access Series of Imaging Studies: Longitudinal MRI data in nondemented and demented older adults. *Journal of Cognitive Neuroscience, 22*(12), 2677-2684. https://doi.org/10.1162/jocn.2009.21407

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701-2715. https://doi.org/10.1162/jocn.2010.21585

Maruyama, T., Takeuchi, H., Taki, Y., Motoki, K., Jeong, H., Kotozaki, Y., Nakagawa, S., Nouchi, R., Iizuka, K., Yokoyama, R., Yamamoto, Y., Hanawa, S., Araki, T., Sakaki, K., Sasaki, Y., Magistro, D., & Kawashima, R. (2018). Effects of Time-Compressed Speech Training on Multiple Functional and Structural Neural Mechanisms Involving the Left Superior Temporal Gyrus. *Neural Plasticity, 2018*, 6574178. https://doi.org/10.1155/2018/6574178

Massaro, D. M. (1987). Categorical partition: A fuzzy-logical model of categorization behavior. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 254-283). Cambridge University Press.

Massaro, D. W., & Oden, G. C. (1980). Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America, 67*(3), 996-1013. https://doi.org/10.1121/1.383941

McAuley, J. D., Henry, M. J., Wedd, A., Pleskac, T. J., & Cesario, J. (2012). Effects of musicality and motivational orientation on auditory category learning: A test of a regulatory-fit hypothesis. *Memory & Cognition, 40*(2), 231-251. https://doi.org/10.3758/s13421-011-0146-4

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86.

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, *60*(1), 1-39. https://doi.org/10.1016/j.cogpsych.2009.06.003

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*, B33-B42. https://doi.org/10.1016/S0010-0277(02)00157-9

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, *60*(1), 65-91. https://doi.org/10.1016/j.jml.2008.07.002

Medin, D. L., & Barsalou, L. W. (1987). Categorization processes and categorical perception. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 455-490). Cambridge University Press.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology*, *64*(2), 199-231. https://doi.org/10.1006/jecp.1996.2343

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS One, 9*(2), e89642. https://doi.org/10.1371/journal.pone.0089642

Myers, E. B., & Blumstein, S. E. (2008). The neural bases of the lexical effect: an fMRI investigation. *Cereb Cortex, 18*(2), 278-288. https://doi.org/10.1093/cercor/bhm053

Myers, E. B., Blumstein, S. E., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science, 20*(7), 895-903. https://doi.org/10.1111/j.1467-9280.2009.02380.x

Myers, E. B., & Swan, K. (2012). Effects of category learning on neural sensitivity to non-native phonetic categories. *Journal of Cognitive Neuroscience, 24*(8), 1695-1708. https://doi.org/10.1162/jocn_a_00243

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425. http://www.ncbi.nlm.nih.gov/pubmed/3615753

Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., Gabrieli, J. D. E., & Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences*, 1-10. https://doi.org/10.1073/pnas.1808412115

Newman, D. A. (2014). Missing Data: Five Practical Guidelines. *Organizational Research Methods, 17*(4), 372-411. https://doi.org/10.1177/1094428114548590

Noordenbos, M. W., & Serniclaes, W. (2015). The categorical perception deficit in dyslexia: A meta-analysis. *Scientific Studies of Reading, 19*(5), 340-359. https://doi.org/10.1080/10888438.2015.1052455

Oden, G. C., & Massaro, D. W. (1978). Integration of featural formation in speech perception. *Psychological Review, 85*, 172-191.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience, 29*(45), 14100-14107. https://doi.org/10.1523/JNEUROSCI.3256-09.2009

Pastore, R. E. (1976). Categorical perception: A critical re-evaluation. In S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, & S. R. Silverman (Eds.), *Hearing and Davis: Essays honoring Hallowell Davis* (pp. 253-264). Washington University Press.

Pastore, R. E., Ahroon, W. A., Baffuto, K. J., Friedman, C., Puleo, J. S., & Fink, E. A. (1977). Common-factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 686-696.

Pavlik Jr, P. I., Hua, H., Williams, J., & Bidelman, G. M. (2013, July 6–9). *Modeling and optimizing forgetting and spacing effects during musical interval training* Proceedings of 6th International Conference on Educational Data Mining, Memphis, TN.

Penhune, V. B., Zatorre, R. J., MacDonald, J. D., & Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex, 6*, 661-672. https://doi.org/10.1093/cercor/6.5.661

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). *Statistical parametric mapping: The analysis of functional brain images*. Elsevier.

Perez-Gay Juarez, F., Sicotte, T., Theriault, C., & Harnad, S. (2019). Category learning can alter perception and its neural correlates. *PLoS One, 14*(12), e0226000. https://doi.org/10.1371/journal.pone.0226000

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America, 130*(1), 461-472. https://doi.org/10.1121/1.3593366

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25*(1-2), 21-52. http://www.ncbi.nlm.nih.gov/pubmed/3581727

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics, 15*(2), 285-290.

Reetzke, R., Xie, Z., Llanos, F., & Chandrasekaran, B. (2018). Tracing the Trajectory of Sensory Plasticity across Different Stages of Speech Learning in Adulthood. *Current Biology, 28*(9), 1419-1427 e1414. https://doi.org/10.1016/j.cub.2018.03.026

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage, 206*, 116189. https://doi.org/10.1016/j.neuroimage.2019.116189

Rosen, S., & Howell, P. (1987). Auditory, articulatory and learning explanations of categorical perception in speech. In S. R. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 113-160). Cambridge University Press.

Ross, B., Jamali, S., & Tremblay, K. L. (2013). Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. *BMC Neuroscience, 14*, 151, Article 151. https://doi.org/10.1186/1471-2202-14-151

Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research, 248*(1-2), 48-59. https://doi.org/10.1016/j.heares.2008.11.012

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(9), 533-536.

Scherg, M., & von Cramon, D. Y. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and clinical Neurophysiology, 65*, 344-360.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*(7), 688-694. https://doi.org/10.1038/nn871

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., Dosch, H. G., Bleeck, S., Stippich, C., & Rupp, A. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience, 8*(9), 1241-1247. https://doi.org/10.1038/nn1530

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication, 41*(1), 71-80. https://doi.org/10.1016/s0167-6393(02)00094-8

Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *Journal of Neuroscience, 34*(33), 10937-10949. https://doi.org/10.1523/jneurosci.5315-13.2014

Seppanen, M., Hamalainen, J., Pesonen, A. K., & Tervaniemi, M. (2012). Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. *Frontiers in Human Neuroscience, 6*, 43. https://doi.org/10.3389/fnhum.2012.00043

Seppanen, M., Hamalainen, J., Pesonen, A. K., & Tervaniemi, M. (2013). Passive sound exposure induces rapid perceptual learning in musicians: event-related potential evidence. *Biological Psychology, 94*(2), 341-353. https://doi.org/10.1016/j.biopsycho.2013.07.004

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience, 23*(13), 5545-5552. https://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., Poldrack, R. A., Bilder, R. M., & Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage, 39*(3), 1064-1080. https://doi.org/10.1016/j.neuroimage.2007.09.031

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research, 25*(2), 547-553. https://doi.org/10.1016/j.cogbrainres.2005.08.007

Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals - musicians can't tell sharp from flat. *Perception & Psychophysics, 21*(5), 399-407. https://doi.org/10.3758/bf03199493

Silva, D. M. R., Rothe-Neves, R., & Melges, D. B. (2020). Long-latency event-related responses to vowels: N1-P2 decomposition by two-step principal component analysis. *International Journal of Psychophysiology*, *148*, 93-102. https://doi.org/10.1016/j.ijpsycho.2019.11.010

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research, 291*, 244-252. https://doi.org/10.1016/j.bbr.2015.05.026

Smits, R., Sereno, J., & Jongman, A. (2006). Categorization of sounds. *Journal of Experimental Psychology: Human Perception and Performance, 32*(3), 733-754. https://doi.org/10.1037/0096-1523.32.3.733

Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(12), E1747-1756. https://doi.org/10.1073/pnas.1523266113

Stevens, K. N. (1981). Constraints imposed by the auditory system on the properties used to classify speech sounds. In T. F. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech*. North Holland.

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomics Bulletin & Review*, *24*(6), 1929-1934. https://doi.org/10.3758/s13423-017-1244-5

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports*, *8*, Article 9223. https://doi.org/10.1038/s41598-018-27571-2

Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., & Kawashima, R. (2011). Effects of training of processing speed on neural systems. *Journal of Neuroscience*, *31*(34), 12139-12148. https://doi.org/10.1523/jneurosci.2948-11.2011

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., & Kawashima, R. (2011). Working memory training using mental calculation impacts regional gray matter of the frontal and parietal regions. *PLoS One*, *6*(8), e23175. https://doi.org/10.1371/journal.pone.0023175

Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(32), 10062-10067. https://doi.org/10.1073/pnas.1505114112

Tong, Y., Melara, R. D., & Rao, A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Research*, *1297*, 80-88. https://doi.org/10.1016/j.brainres.2009.07.089

Toscano, J. C., Anderson, N. D., Fabiani, M., Gratton, G., & Garnsey, S. M. (2018). The time-course of cortical responses to speech revealed by fast optical imaging. *Brain and Language, 184*, 32-42. https://doi.org/10.1016/j.bandl.2018.06.006

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization: electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science, 21*(10), 1532-1540. https://doi.org/10.1177/0956797610384142

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and Hearing, 22*(2), 79-90. https://doi.org/10.1097/00003446-200104000-00001

Tremblay, K. L., Inoue, K., McClannahan, K., & Ross, B. (2010). Repeated stimulus exposure alters the way sound is encoded in the human brain. *PLoS One, 5*(4), e10283. https://doi.org/10.1371/journal.pone.0010283

Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research, 45*(3), 564-572. https://doi.org/10.1044/1092-4388(2002/045)

Tremblay, K. L., Ross, B., Inoue, K., McClannahan, K., & Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? *Frontiers in Systems Neuroscience, 8*(28), 1-13. https://doi.org/10.3389/fnsys.2014.00028

Tremblay, K. L., Shahin, A. J., Picton, T., & Ross, B. (2009). Auditory training alters the physiological detection of stimulus-specific cues in humans. *Clinical Neurophysiology, 120*(1), 128-135. https://doi.org/10.1016/j.clinph.2008.10.005

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences, 63*, 87-93. https://doi.org/10.1016/j.paid.2014.01.057

Wang, X. Y., Wang, S. P., Fan, Y. B., Huang, D., & Zhang, Y. (2017). Speech-specific categorical perception deficit in autism: An Event-Related Potential study of lexical tone processing in Mandarin-speaking children. *Scientific Reports, 7*, 1-12, Article 43254. https://doi.org/10.1038/srep43254

Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., & Schneider, P. (2014). Increased volume and function of right auditory cortex as a marker for absolute pitch. *Cerebral Cortex, 24*(5), 1127-1137. https://doi.org/10.1093/cercor/bhs391

Wenger, E., Papadaki, E., Werner, A., Kühn, S., & Lindenberger, U. (2020). Observing plasticity of the auditory system: Volumetric decreases along with increased functional connectivity in aspiring professional musicians. *bioRxiv [preprint]*. https://doi.org/10.1101/2020.11.24.393751

Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology, 41*(1), 48-61. http://www.ncbi.nlm.nih.gov/pubmed/3502888

West, B. T., Welch, K. B., & Gałecki, A. T. (2015). *Linear mixed models: A practical guide using statistical software* (2 ed.). CRC Press.

Westbury, C. F., Zatorre, R. J., & Evans, A. C. (1999). Quantifying variability in the planum temporale: A probability map. *Cerebral Cortex, 9*, 392-405. https://doi.org/10.1093/cercor/9.4.392

Whitwell, J. L., Crum, W. R., Watt, H. C., & Fox, N. C. (2001). Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *AJNR. American Journal of Neuroradiology, 22*(8), 1483-1489.

Wisniewski, M. G., Ball, N. J., Zakrzewski, A. C., Iyer, N., Thompson, E. R., & Spencer, N. (2020). Auditory detection learning is accompanied by plasticity in the auditory evoked potential. *Neuroscience Letters, 721*, Article 134781. https://doi.org/10.1016/j.neulet.2020.134781

Wong, P. C., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping, 28*(10), 995-1006. https://doi.org/10.1002/hbm.20330

Wong, P. C., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's Gyrus and linguistic pitch learning. *Cerebral Cortex, 18*(4), 828-836. https://doi.org/10.1093/cercor/bhm115

Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers. *Frontiers in Psychology, 6*, 436. https://doi.org/10.3389/fpsyg.2015.00436

Yi, H. G., & Chandrasekaran, B. (2016). Auditory categories with separable decision boundaries are learned faster with full feedback than with minimal feedback. *Journal of the Acoustic Society of America, 140*(2), 1332. https://doi.org/10.1121/1.4961163

Yoo, J., & Bidelman, G. M. (2019). Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hearing Research, 377*, 189-195. https://doi.org/https://doi.org/10.1016/j.heares.2019.03.021

Zatorre, R., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience, 15*(4), 528-536. https://doi.org/10.1038/nn.3045

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science (New York, N.Y.), 256*(5058), 846-849. https://doi.org/10.1126/science.1589767

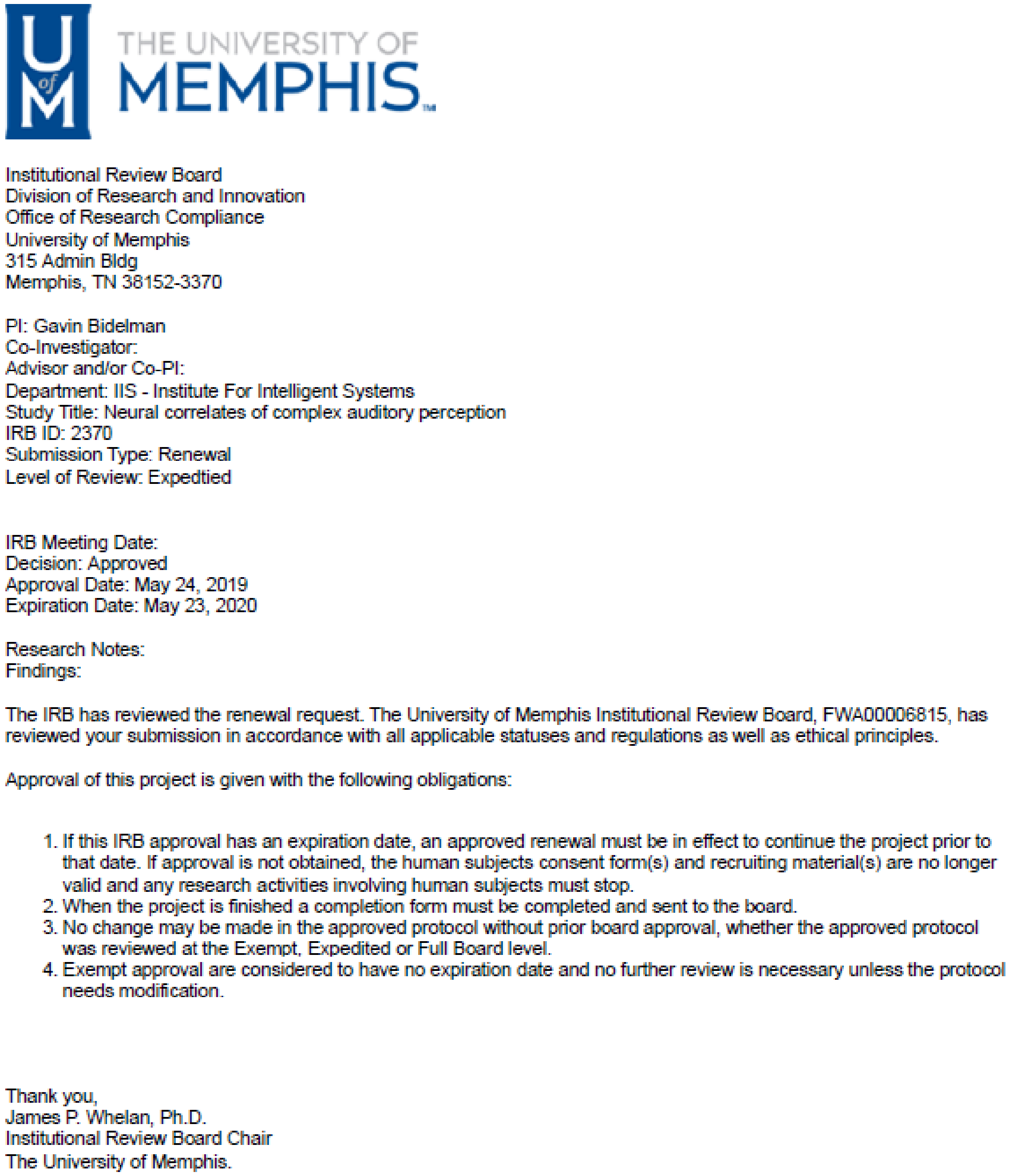
Zatorre, R. J., & Halpern, A. R. (1979). Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception and Psychophysics*, *26*(5), 384-395.

Zendel, B. R., Tremblay, C. D., Belleville, S., & Peretz, I. (2015). The impact of musicianship on the cortical mechanisms related to separating speech from background noise. *Journal of Cognitive Neuroscience*, *27*(5), 1044-1059. https://doi.org/10.1162/jocn_a_00758

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *Neuroimage*, *26*(3), 703-720. https://doi.org/10.1016/j.neuroimage.2005.02.040

Appendix

IRB Approval: Protocol #2370



Institutional Review Board
Division of Research and Innovation
Office of Research Compliance
University of Memphis
315 Admin Bldg
Memphis, TN 38152-3370

PI: Gavin Bidelman
Co-Investigator:
Advisor and/or Co-PI:
Department: IIS - Institute For Intelligent Systems
Study Title: Neural correlates of complex auditory perception
IRB ID: 2370
Submission Type: Renewal
Level of Review: Expedtied

IRB Meeting Date:
Decision: Approved
Approval Date: May 24, 2019
Expiration Date: May 23, 2020

Research Notes:
Findings:

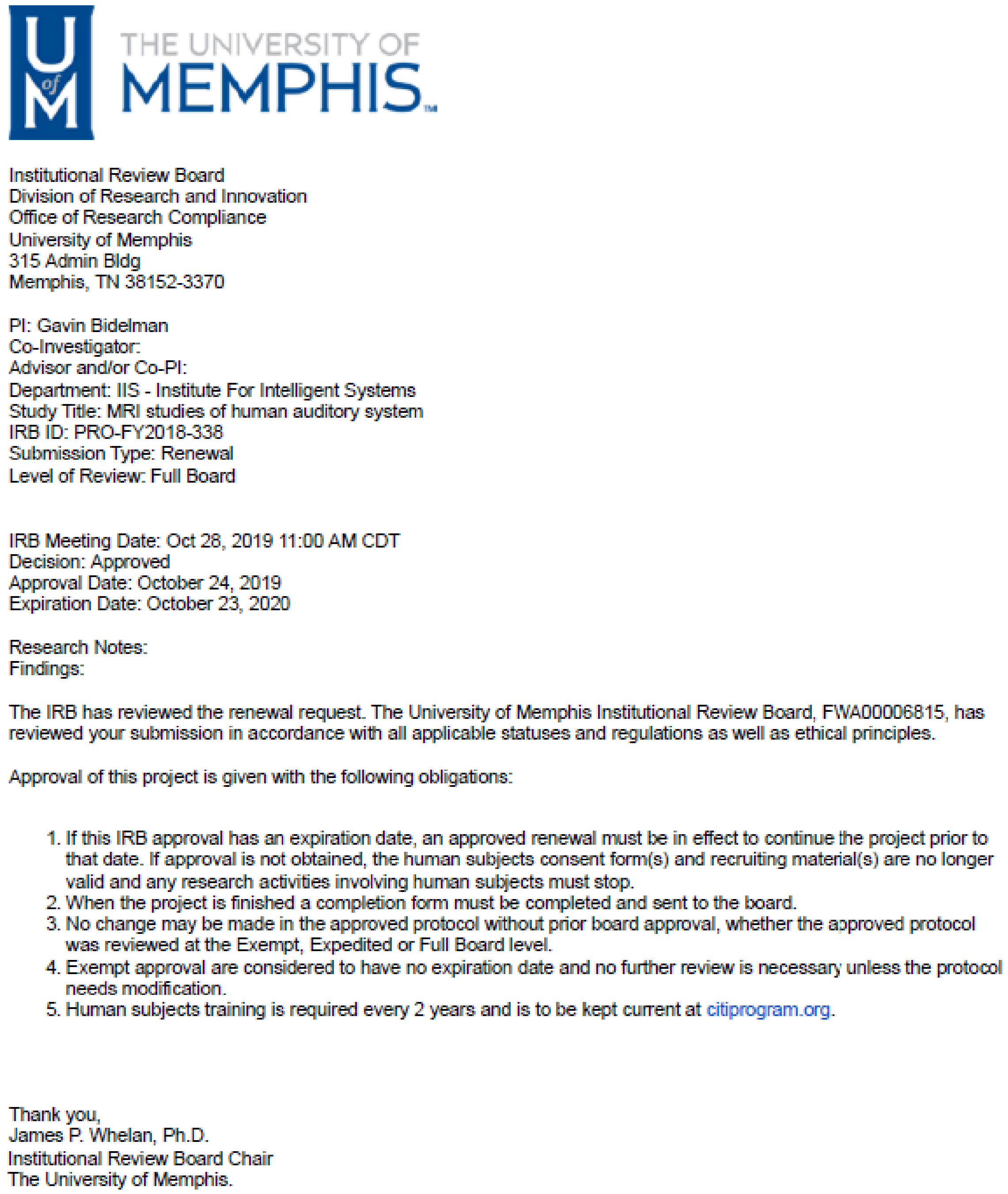
The IRB has reviewed the renewal request. The University of Memphis Institutional Review Board, FWA00006815, has reviewed your submission in accordance with all applicable statuses and regulations as well as ethical principles.

Approval of this project is given with the following obligations:

1. If this IRB approval has an expiration date, an approved renewal must be in effect to continue the project prior to that date. If approval is not obtained, the human subjects consent form(s) and recruiting material(s) are no longer valid and any research activities involving human subjects must stop.
2. When the project is finished a completion form must be completed and sent to the board.
3. No change may be made in the approved protocol without prior board approval, whether the approved protocol was reviewed at the Exempt, Expedited or Full Board level.
4. Exempt approval are considered to have no expiration date and no further review is necessary unless the protocol needs modification.

Thank you,
James P. Whelan, Ph.D.
Institutional Review Board Chair
The University of Memphis.

IRB Approval: Protocol #PRO-FY2018-338



THE UNIVERSITY OF MEMPHIS.

Institutional Review Board
Division of Research and Innovation
Office of Research Compliance
University of Memphis
315 Admin Bldg
Memphis, TN 38152-3370

PI: Gavin Bidelman
Co-Investigator:
Advisor and/or Co-PI:
Department: IIS - Institute For Intelligent Systems
Study Title: MRI studies of human auditory system
IRB ID: PRO-FY2018-338
Submission Type: Renewal
Level of Review: Full Board

IRB Meeting Date: Oct 28, 2019 11:00 AM CDT
Decision: Approved
Approval Date: October 24, 2019
Expiration Date: October 23, 2020

Research Notes:
Findings:

The IRB has reviewed the renewal request. The University of Memphis Institutional Review Board, FWA00006815, has reviewed your submission in accordance with all applicable statuses and regulations as well as ethical principles.

Approval of this project is given with the following obligations:

1. If this IRB approval has an expiration date, an approved renewal must be in effect to continue the project prior to that date. If approval is not obtained, the human subjects consent form(s) and recruiting material(s) are no longer valid and any research activities involving human subjects must stop.
2. When the project is finished a completion form must be completed and sent to the board.
3. No change may be made in the approved protocol without prior board approval, whether the approved protocol was reviewed at the Exempt, Expedited or Full Board level.
4. Exempt approval are considered to have no expiration date and no further review is necessary unless the protocol needs modification.
5. Human subjects training is required every 2 years and is to be kept current at citiprogram.org.

Thank you,
James P. Whelan, Ph.D.
Institutional Review Board Chair
The University of Memphis.