Music
RJ Zatorre, McGill University, Montreal, QC, Canada
AR Halpern, Bucknell University, Lewisburg, PA, USA
© 2015 Elsevier Inc. All rights reserved.

Introduction
Understanding the neural bases of music provides valuable insights into some of our most complex human cognitive abilities. Musical perceptual skills are present from birth, but the ability to produce instrumental music requires training. The training variable is a useful one to study plasticity of this complex system. Memory for music shows some paradoxes: people have vast mental repositories of familiar music that can be accessed quickly and that conserve important musical properties in the form of auditory imagery. However, learning new music is slow and effortful. Finally, music not only involves cognitive mechanisms but also almost by definition captures and engenders emotional response. In this article, we present neuroimaging evidence of the neural underpinnings of perception and production mechanisms, memory (including auditory imagery), musical emotion, and the relationships between neural mediation of music and lifelong and short-term training.

Perception/Production
A global sketch of the auditory nervous system would start with transduction of sound into action potentials by peripheral mechanisms. A number of brain stem nuclei are then involved in representing stimulus features, with information thence flowing to core areas of the auditory cortex, located in the medial portion of Heschl’s gyrus (HG), and surrounding belt and parabelt areas in the superior temporal gyrus (STG). Two pathways emerge at this point, a more ventral one that integrates auditory and visual information and connects to medial temporal memory structures and a more dorsal one that connects to parietal and premotor regions important for sensory–motor interactions (Figure 1). These pathways then target distinct frontal-lobe regions, forming bidirectional processing loops (for a review, see Rauschecker and Scott (2009)).

Electrical measures of frequency coding in brain stem structures indicate that critical aspects of music, including pitch salience, consonance and dissonance, and harmonic relationships, emerge already at this early stage of processing (Bidelman, 2013). These measures reflect relevant properties of the stimulus but may also be dependent on cortical efferent influences. At the cortical level, apart from the well-established principle of tonotopy (Saenz & Langers, 2014), fMRI studies have detailed single-voxel organization that reflects sensitivity to spectral and temporal modulations (Schönwiesner & Zatorre, 2009) as well as coding of harmonic and octave relations (Moerel et al., 2013), features likely of special significance for music. Several fMRI studies point to lateral HG or adjacent areas as regions where pitch information is encoded (Griffiths & Hall, 2012; Norman-Haignere, Kanwisher, & Mcdermott, 2013). As pitches are combined to form melodies, cortical areas beyond these pitch-sensitive regions come into play, with a right-sided predominance (Hyde, Peretz, & Zatorre, 2008; Patterson et al., 2002). The relative asymmetry for processing of pitch patterns, long noted in lesion studies (Stewart et al., 2006), has been attributed to enhanced spectral resolution in the right auditory cortex (Schönwiesener, Rubsamen, & Von Cramon, 2005; Zatorre & Gandour, 2007). Abstract coding of relationships across pitches has recently been demonstrated in both ventral and dorsal auditory areas using multivariate pattern analysis for musical interval categories (Klein & Zatorre, in press) and for melodic contour (Lee et al., 2011), two features identified as critical for melody perception in cognitive studies (Dowling & Harwood, 1986).

Since sounds exist only in time, computing relationships between tones in a melody requires a working memory mechanism to maintain information. Frontal–STG interactions via the aforementioned loops are important in this function as shown by positron emission tomography (PET) (Zatorre, Evans and Meyer 1994), fMRI (Koelsch et al., 2009), and MEG (Grimault et al., 2014) findings, in accord with working memory models from other domains. Disruptions of frontotemporal connections are associated with tonal perceptual disorders (Hyde, Zatorre, & Peretz, 2011), as evidenced by both structural imaging (Loul, Alsop, & Schlaug, 2009) and functional imaging (Albouy et al., 2013). Working memory is also important for manipulating information: recognizing a melody as the same when transformed via temporal reversal or via pitch transposition recruits the intraparietal sulcus (Foster, Halpern, & Zatorre, 2013), a key dorsal pathway also implicated in visuomotor transformations.

Listening to a melody is an active process because it involves expectancies that are generated based upon a listener’s implicit knowledge of musical rules (Pearce & Wiggins, 2006). Unexpected violations of contour (Tervaniemi et al., 2001), interval size (Trainor, Mcdonald, & Alain, 2002), or harmonies (Koelsch, Gunter, & Friederici, 2000) elicit physiological responses that likely arise from frontotemporal interactions (Tillmann et al., 2006). Similarly, metrical structure provides a framework for expectancies about when events will occur (Vuust et al., 2009) and involve functional interactions between the superior temporal and premotor cortices (Chen, Zatorre, & Pennhune, 2006), as well as with the basal ganglia (Grahn & Rowe, 2013; Kung et al., 2013). These interactions with motor-related neural systems may provide a substrate for the link between temporal structure in music and movement, including dance (Calvo-Merino et al., 2005).

Movement of course is needed to produce sound, and rapid and highly controlled actions are required to make music. When trained musicians are scanned while performing on a keyboard, dorsal-stream networks are recruited, including the premotor and supplementary motor cortices, as well as the...
intraparietal region (Bengtsson & Ullén, 2006; Brown et al., 2013) reflecting online sensory–motor adjustments via interactions with auditory areas. Integration of auditory and motor systems is indicated by consistent evidence of activity of motor-related networks for perceiving musical patterns that musicians know how to play (Bangert et al., 2006; Baumann et al., 2005). The premotor cortex is especially important in enabling these perceptuomotor interactions (for review, see Zatorre, Chen, and Penhune, 2007). Studies of the neural control of the vocal system for singing provide valuable comparison with other systems (for review, see Zarate, 2013). Activation sites overlap with the vocal motor system for speech, but with more right-sided laterality (Ozdemir, Norton, & Schlaug, 2006), likely related to the higher degree of pitch control required to sing (Zatorre & Baum, 2012). A specific role for the anterior insula in singing has also been suggested based on its role in integration of somatosensory and proprioceptive information (Kleber et al., 2013), sources of feedback that become more important as expertise is developed (Kleber et al., 2010).

Memory

How well does the brain handle memory for melodies? The answer depends on the type of memory: episodic memory (which tunes did I just hear?) is surprisingly poor, especially when contrasted with memory for lists of words, pictures, or faces. On the other hand, semantic memory (is that tune one I have learned in my life?) is quite good (Halpern & Bartlett, 2010). Even nonmusicians have a large repertory of familiar tunes, maintain it over the life span (Rubin, Rahhal, & Poon, 1998), and can maintain at least an approximate memory of absolute pitch (Schellenberg & Trehub, 2003) and tempo (Levitin & Cook, 1996). Although ERP responses to a wrong terminal note for familiar tunes are stronger in musicians than nonmusicians, the latter do show a reliable late positive reaction to incorrect notes (Besson & Faita, 1995).

Brain mapping studies in musical memory are not very abundant for either type of memory. In episodic memory, most studies have considered recognition (recall is rarely
used in music memory paradigms). Studies show different networks active depending on familiarity of the materials and also how the recognition responses are gathered. One PET study (Platel et al., 2003) showed different patterns of activation during episodic recognition of familiar nonverbal tunes (largely right-sided activation, including the precuneus and superior frontal gyrus) versus unfamiliar tunes (various bilateral frontal regions). Using more sensitive event-related fMRI trials can be partitioned into hits and correct rejections; more successful recognition (hit-CRs) of unfamiliar tunes is correlated with activity in the right hippocampus (Watanabe, Yagishita, & Kikyo, 2008). Confidence measures gathered on recognition reveal that the right posterior parietal area is implicated in the high-confidence hits (vs. CRs) for unfamiliar tunes (Klostermann, Loui, & Shimamura, 2009).

Semantic memory engages yet different networks, with more bilateral activation. When asked whether a tune is familiar from life, or was newly composed, the same PET study referred to earlier showed a largely left-lateralized network including the frontal and temporal areas and angular gyrus (Platel et al., 2003) although Peretz et al. (2009), using a passive task with similar materials, additionally implicated superior temporal regions, right more than left. For songs (melody plus words), it seems that separate networks mediate the semantic lookup for melodies (right middle temporal and bilateral posterior temporal) and words (the left fusiform gyrus and the left inferior occipital gyrus), with the left inferior temporal cortex mediating the integrated representation (Saito et al., 2012). fMRI data also show that the medial prefrontal cortex, an area that supports self-referential processes, is recruited when music triggers autobiographical memories, thus linking this system to the strong emotional memories that music can evoke (Janata, 2009).

Thus, we see that similar to results from music perception tasks, episodic memory, particularly for robust memory experiences, seems to depend on a right-lateralized network of frontal and parietal areas. However, semantic memory for music may engage fewer music-specific and more general memory representation systems, given how many associations people have to well-known music: *White Christmas* is likely tied to memories of one’s own holidays as well as one’s general knowledge of the movie it appeared in or opinions of Bing Crosby as a singer.

**Imagery**

One special kind of memory for music is musical imagery, or the phenomenology of “hearing” a tune playing in the mind. This experience can be indexed behaviorally: even nonmusicians can make consistent and accurate judgment of mental tempo or pitch (Halpern, 1988; Janata, 2012) (for review, see Hubbard, 2010). These internal representations are sufficiently precise to elicit a response from the auditory cortex if there is a mismatch between imagined and real tones (Herholz et al., 2008).

Given the perceptual-like nature of auditory imagery, neuroimaging studies have investigated the extent to which the brain co-opts perceptual networks during imagery tasks. Interestingly, extant studies provide no convincing evidence for activation of the primary auditory cortex during auditory imagery (Zatorre & Halpern, 2005), suggesting one way the brain keeps track of the difference between imagery, perceived as internal, and hallucinations, which do involve the primary cortices (Kompus et al., 2013).

Perceptual systems are involved at the level of the secondary auditory cortex (Figure 2). This response may be bilateral for imagery of verbal music, but is often right-lateralized, consistent with music perception, particularly when nonverbal components such as timbre are imagined (Halpern et al., 2004) or the musical material needs to be manipulated, such as mental tune reversal (Zatorre, Halpern, & Bouffard, 2010). Consistent with this conclusion, a lesion study (Zatorre & Halpern, 1993) showed that right, but not left, anterior temporal lobectomy disturbed an imagined pitch comparison task.

Beyond perceptual representation, auditory imagery involves sequencing, memory, and internal attention. Consistent with the sequencing aspect, the supplementary (or pre-supplementary) motor cortex shows up frequently as an activation site in a variety of auditory imagery tasks, such as anticipation of the next tune in a list (Leaver et al., 2009), mental pitch comparison (Zatorre et al., 1996), and mentally

---

**Figure 2**  
Functional MRI activation patterns associated with musical imagery. Areas recruited by listening or imagining a musical passage are shown in orange and blue, respectively. Areas of overlap between the two tasks are shown in green. The auditory cortices in the superior temporal gyrus are extensively recruited when hearing music; regions that respond to sound but fall outside the primary auditory cortex are also recruited in silence during musical imagery. Premotor and supplementary motor regions are also recruited in common. The inferior frontal cortices are active only during imagery, reflecting perhaps retrieval functions. Reproduced from Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, 24, 1382–1397.

---

Emotion

Music’s ability to elicit a wide range of emotional responses provides a challenge to psychologists and emotion theorists. Some of the many circumstances leading to music-induced emotion have been codified and analyzed (Juslin & Sloboda, 2010), but many questions remain as to how musical emotions fit within a broader theoretical framework (Zentner, Grandjean, & Scherer, 2008). Nonetheless, neuroimaging has contributed crucial empirical information that helps to constrain these models and link musical emotions to neural systems associated with general affective processing mechanisms (Koelsch, 2010).

Functional imaging studies have shown that dissonant stimuli, which are perceived as unpleasant, elicit reciprocal changes in paralimbic and neocortical areas. Increasing dissonance is associated with reductions of activity in orbital and ventromedial prefrontal areas (Blood et al., 1999) in accord with studies of valence using other stimuli (Kringelbach & Rolls, 2004), while recruitment is seen in parahippocampal regions, which aligns well with the role of this structure from lesion studies (Gosselin et al., 2006). Dissonance also decreases activity within the amygdala (Koelsch et al., 2006). More complex emotions elicited by music (joy, nostalgia, and sadness) induce modulation of a similar network of structures (Trost et al., 2012), which can largely be explained on the basis of two orthogonal dimensions: arousal versus valence.

Turning to positive effect, peak positive emotion can be indexed by the presence of ‘chills,’ a physiologically identifiable moment of autonomic nervous system arousal that is experienced as highly pleasurable (Grewe et al., 2007; Salimpoor et al., 2009). Brain activity associated with musical chills is observed in a number of brain circuits associated with emotion and arousal and particularly in the striatum (Blood & Zatorre, 2001), a component of the mesolimbic reward system that responds consistently to primary and secondary rewards, such as food, money, and erotic stimuli (Sescousse et al., 2013). The striatal response also occurs in the absence of chills to pleasurable music (Mitterschiffthaler et al., 2007) and is functionally linked to midbrain regions (Menon & Levitin, 2005), supporting its role in a dopaminergic circuit. Direct evidence for involvement of dopamine was provided via PET-raclopride measures (Salimpoor et al., 2011) showing significant binding of dopamine in both the dorsal striatum and the ventral striatum during highly pleasurable music. The activity in these two structures was dissociated via fMRI in terms of the anticipation versus experience of peak pleasure, which were maximal in the dorsal and ventral striatum, respectively (Figure 3). This finding can be linked to reward prediction (O’doherty, 2011) and models of temporal difference learning (Zald & Zatorre, 2011), as well as to music-theoretical analyses of expectancies (Huron, 2006). Outside the context of peak pleasure, monetary valuation of music is also predicted by ventral striatum activity even for novel music (Salimpoor et al., 2009), despite the absence of actual motor execution in these tasks. Intraparietal sulcus is active during manipulation of imagined sounds, like mental tune reversal or transposition (Foster et al., 2013); and individuals who report having more vivid auditory imagery show more (and more right-lateralized) activity in this structure (Zatorre et al., 2010), as well as in anterior STG (Herholz et al., 2012).

Finally, as imagery is closely linked to memory, we should not be surprised that working memory areas such as the dorsolateral prefrontal cortex are active during imagery. Not only this area is active during melody retrieval, but also its activity is functionally coupled to the right secondary auditory area (Herholz, Halpern, & Zatorre, 2012). Furthermore, that relationship is stronger among people with higher self-reported imagery. Thus, we see that brain mapping relationships reflect the interaction of perceptual, memory, and control areas needed to generate and manipulate tunes in the mind. | Author's personal copy

Figure 3 Neural correlates of processing pleasurable music. (a) Spatial conjunction analysis between (11C) raclopride positron emission tomography, sensitive to dopamine release, and fMRI while listeners heard pleasurable music. Increased hemodynamic activity in dopaminergic regions was elicited in the ventral striatum (VS) during peak emotional moments (marked by ‘chills’) and in the dorsal striatum (DS) during the anticipatory phase, preceding chills. Reproduced from Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nature Neuroscience, 14, 257–264. (b) fMRI scanning showing that the best predictor of reward value of new music (as marked by monetary bids in an auction paradigm) was activity in the ventral striatum, particularly the nucleus accumbens (NAcc), which also showed increased functional connectivity with the superior temporal gyrus (STG) and the right inferior frontal gyrus (IFG) as musical stimuli gained reward value. Reproduced from Salimpoor, V. N., Van Den Bosch, I., Kovacevic, N., et al. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. Science, 340, 216–219.
et al., 2013). Moreover, as reward value increases, functional coupling between the ventral striatum and the amygdala, ventromedial frontal cortex, and auditory cortex also increases. The finding of interactions between reward system and auditory areas suggests that the formation of predictions and expectancies based on abstract knowledge of musical patterns, as instantiated via the auditory frontal loops described earlier, may play a critical role in the way that musical pleasure arises.

**Plasticity**

Most of the brain networks described earlier are subject to experience-dependent changes; indeed, musical training has been highlighted as a framework for understanding plasticity in a cognitive neuroscience context (Herholz & Zatorre, 2012). Cross-sectional studies comparing groups of musically trained and untrained individuals have consistently documented both functional and structural differences in a number of systems (Wan & Schlaug, 2010). At the brain stem level, musicians demonstrate faster and more accurate encoding of periodic auditory stimuli, which in turn is associated with enhancements in capacities such as hearing in noise, timing, or binaural processing (Strait & Kraus, 2014). At the cortical level, musicians show enhanced evoked responses to tones (Pantev et al., 2003), which are related to morphological changes in the auditory cortex and which correlate with behavior on auditory perceptual tasks (Foster & Zatorre, 2010; Schneider et al., 2002). Anatomical features that characterize musicians include not only greater volume or thickness of auditory cortices but also changes in the frontal and motor cortices (Bermudez et al., 2009; Gaser & Schlaug, 2003); cortical anatomy in the intraparietal regions is specifically predictive of melody transposition performance (Foster & Zatorre, 2010), in keeping with the functional role of the dorsal pathway. White-matter enhancements have also been documented in several tracts, including the descending motor pathways of pianists as a function of amount of practice they engaged in during childhood (Bengtsson et al., 2005). Greater interhemispheric connectivity in musicians is suggested by findings of greater white-matter volume in the anterior corpus callosum (Schlaug et al., 1995). Diffusion measures of white-matter organization also implicate the portion of the corpus callosum that interconnects the left and right motor/premotor cortices in musicians whose training started early in life (Steele et al., 2013). The anatomical connectivity between auditory areas and frontal cortices is also greater with musical training (Halwani et al., 2011), which presumably is related to the enhanced sensory–motor and working memory skills that depend on these circuits, as described earlier.

Cross-sectional studies cannot distinguish causal effects from preexisting differences. Longitudinal or controlled training studies provide evidence that both factors may play a role. At the anatomical level, Hyde et al. (2009) showed that 15 months of music training in school-aged children resulted in morphological changes in the auditory and motor cortices, as well as the corpus callosum, supporting the conclusions from group comparisons. Functionally, there is also evidence that musical training in children enhances cortical responses to tones (Fujioka et al., 2006). Plasticity can also be seen in adults: explicit piano training results in enhanced coupling between auditory and motor systems (Lahav, Saltzman, & Schlaug, 2007), even after a single session of learning (Chen, Rae, & Watkins, 2012); this type of auditory motor enhancement requires explicit motor training and does not emerge with passive exposure (Lappe et al., 2008). Perceptual training does however result in changes to auditory cortical responses measured both with MEG (Bosnyak, Eaton, & Roberts, 2004) and with fMRI (Caab, Gaser, & Schlaug, 2006). The clear evidence for experience-dependent plasticity does not necessarily preclude the existence of predispositions. Indeed, individual differences in learning rate or attainment are frequently noted, and a number of studies have identified both functional and structural features of sensory–motor networks that are predictive of how well individuals will learn (for review, see Zatorre (2013)).

**See also:** INTRODUCTION TO COGNITIVE NEUROSCIENCE: Prediction and Expectation; Working Memory; INTRODUCTION TO SYSTEMS: Reward; Visuomotor Integration.

**References**


Klein, M. E., & Zatorre, R. J. (in press). Representations of invariant, categorically perceived classes of musical sounds are decodable by multivariate pattern analysis of locally distributed BOLD responses in superior temporal and intraparietal sulci. *Cerebral Cortex*.


