Playing Music for a Smarter Ear: Cognitive, Perceptual and Neurobiological Evidence

Dana Strait and Nina Kraus
Northwestern University

Human hearing depends on a combination of cognitive and sensory processes that function by means of an interactive circuitry of bottom-up and top-down neural pathways, extending from the cochlea to the cortex and back again. Given that similar neural pathways are recruited to process sounds related to both music and language, it is not surprising that the auditory expertise gained over years of consistent music practice fine-tunes the human auditory system in a comprehensive fashion, strengthening neurobiological and cognitive underpinnings of both music and speech processing. Given this, we argue not only that common neural mechanisms for speech and music exist, but that experience in music leads to enhancements in sensory and cognitive contributors to speech processing. Of specific interest is the potential for music training to bolster neural mechanisms that undergird language-related skills, such as reading and hearing speech in background noise, which are critical to academic progress, emotional health, and vocational success.

Received December 12, 2010, accepted July 21, 2011.

Key words: musicians, language, attention, memory, brain

Debates concerning music training’s impact on general cognitive and perceptual abilities are easily sparked, partly due to the widespread popular interest that has stemmed from poorly controlled research. Effects of musical experience have been proposed through comparing musician and nonmusician groups without first ensuring that these groups do not differ according to overarching factors such as IQ, socioeconomic status, and level of education, to name a few (Schellenberg, 2005). Even amidst this lack of scientific control, compelling evidence has arisen to support the power of music practice to shape basic sensory and cognitive auditory function (Kraus & Chandrasekaran, 2010). In addition to contributing to great amusement and well-being, practicing music does, in fact, appear to make you smarter – at least, smarter when it comes to how you hear. In this review, we provide neurobiological evidence that music training shapes human auditory function not only as it relates to music, but also as it relates to speech and other language-related abilities, such as reading. This evidence is presented in the context of discussions of common neural mechanisms for processing speech and music.

Music Training Makes us Better Listeners: No Easy Feat

Although cortical specializations for music and speech have each been established (Abrams et al., 2010; Brown, Martinez, & Parsons, 2006; Rogalsky, Rong, Saberi, & Hickok, 2011; Zatorre, Belin, & Penhune, 2002), there is no doubt that the human brain also recruits similar cortical mechanisms for processing sound in both domains (Koelsch et al., 2002; Patel, 2003; Rogalsky et al., 2011; Zatorre & Gandour, 2008). Research has substantiated direct links between musicianship and human sound processing, both within and outside of the domain of music. This work suggests that musicians are “better listeners” than nonmusicians with regard to how they perceive and neurally process sound in any domain. From a neurobiological standpoint, that music training has the power to make people better listeners is no simple feat. Here, we review evidence for musicians’ auditory processing enhancements in the context of a brief description of their underlying neuronal mechanisms, emphasizing the potential of music training to strengthen cognitive control over sensory function.

Perceptually, musicians demonstrate heightened auditory acuity, as evidenced by their enhanced ability (compared to nonmusicians) to discriminate pitch discrepancies and temporal gaps between sounds (Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Micheyl, Delhomméau, Perrot, & Oxenham, 2006; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011; Strait, Kraus, Parbery-Clark, & Ashley, 2010). Musicians’ fine-tuned auditory perception may account for their increased
sensitivity to the pitch and temporal components of language and music and their enhanced cortical evoked potentials to deviations in the pitch and meter of a sound stream (Chobert, Marie, Francois, Schön & Besson, 2011; Marie, Magne, & Besson, 2010; Marques, Moreno, Castro, & Besson, 2007; Schön, Magne, & Besson, 2004; Tervaniemi, Ilvonen, Karma, Alho, & Naatanen, 1997; van Zuijjen, Sussman, Winkler, Naatanen, & Tervaniemi, 2005). Even subcortically, musicians demonstrate faster and more robust auditory brainstem responses to music (Lee, Skoe, Kraus, & Ashley, 2009; Musacchia, Sams, Skoe, & Kraus, 2007), speech (Bideman, Gandour, & Krishnan, 2009; Bideman & Krishnan, 2010; Musacchia et al., 2007; Parbery-Clark, Skoe, & Kraus, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007), and emotional communication sounds (Strait, Kraus, Skoe, & Ashley, 2009).

Musicians’ subcortical enhancements do not reflect a simple “volume knob” effect, resulting in overall larger and faster responses. Rather, musicians demonstrate selective enhancements for the most behaviorally relevant aspects of sound (e.g., the upper note of a musical interval, which often carries the melody; Figure 1d), a finding that is consistent with animal work indicating selective neural tuning to behaviorally relevant stimuli (Fritz, Elhilali, & Shamma, 2007; Suga, Xiao, Ma, & Ji, 2002; Woolley, Gill, & Theunissen, 2006). In fact, the ability of the nervous system to modify auditory function based on survival-promoting interactions with its environment is a fundamental operating principle seen throughout the animal kingdom. Selective tuning is likely dependent, at least in part, on cortico-subcortico reciprocity facilitated by activation of the neuromodulatory system (notably, acetylcholine release from the basal forebrain; Ji & Suga, 2009; Suga & Ma, 2003).

Even beyond subcortical and cortical functional enhancements, music training may shape auditory function in structures as peripheral as the cochlea, with musicians demonstrating a greater degree of efferent control over outer hair cell activity along the basilar membrane than nonmusicians (Brashears, Morlet, Berlin, & Hood, 2003; Perrot, Micheyl, Khalfa, & Collet, 1999). Such comprehensive perceptual and neural enhancements may be driven, at least in part, by strengthened cognitive control over basic auditory processing, as engendered by auditory attention (Strait et al., 2010; Tervaniemi et al., 2009) and working memory (Ho, Cheung, & Chan, 2003; Pallesen et al., 2010; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark et al., 2011) – two auditory cognitive skills that are enhanced in musicians (Figure 1b, c). In fact, it has been repeatedly demonstrated that how we hear (e.g., how well we are able to hear speech in noisy environments, discriminate frequencies, or separate rapidly occurring sounds temporally) cannot be predicted by hearing thresholds alone (He, Dubno, & Mills, 1998; Killion & Niquette, 2000). Auditory acuity can be predicted by prefrontal cortical activation (Wong, Ettlinger, Sheppard, Gunasekera, & Dhar, 2010) and is not only shaped by cognitive control over auditory processing (Conway, Cowan, & Bunting, 2001; Hafter, Sarampalis, & Loui, 2008; Strait et al., 2010) but also by the languages we speak (Bent, Bradlow, & Wright, 2006; Krishnan, Gandour, & Bidelman, 2010; Krishnan, Xu, Gandour, & Cariani, 2005; So & Best, 2010; Van Engen & Bradlow, 2007), the activities we perform (Marques et al., 2007; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Schön et al., 2004), and auditory training (Alain, Campeau, & Tremblay, 2010; Song, Skoe, Wong, & Kraus, 2008; Tremblay, Shahin, Picton, & Ross, 2009; Wright & Zhang, 2009).

Hearing does not solely rely on an auditory relay system that passes sound from the cochlea to the cortex. As described above, it is influenced by cognitive function, behavioral significance, and experience. Such influences are facilitated by the human auditory system’s complex symphony of afferent (i.e., bottom-up) and efferent (i.e., top-down) neuronal pathways that interact to shape how we experience our auditory world (Figure 1a). That the brain responds to sound in a bottom-up manner is intuitive: sound waves are converted into mechanical vibrations followed by their transduction along the basilar membrane into a neural code. This neural code then gets passed along in quasi-relay fashion, up the auditory pathway from the cochlea to the cortex. A more holistic and perhaps less intuitive view of the human auditory system must consider how auditory function is guided both by the languages we speak and by the neural activity. The human top-down auditory system is both vast and intricate, matching the volume of ascending fibers (Winer, 2005) and demonstrating more complex interconnectivity with subcortical processing sites than other sensory domains, such as vision. This is due to the sheer number of innervated nuclei between the cochlea and the cortex and the alternative pathways by which signals can travel from one nucleus to the next. Furthermore, efferent neuronal paths originate not only in the auditory cortex but also in non-auditory areas such as the limbic system and cognitive centers of memory and attention. By means of these pathways, the efferent network changes basic response properties at the aforementioned auditory relay sites through top-down modulation.

A central contribution of this top-down system to auditory function is to improve sound processing at low-level
sensory centers by tuning them to relevant auditory input (Bajo, Nodal, Moore, & King, 2010; Suga, 2008). The strength of the efferent system predicts how well humans hear in challenging listening environments (e.g., in background noise), and individuals with stronger efferent control over auditory processing are more apt to improve at an auditory task with brief experience (de Boer & Thornton, 2008). This bodes well for musicians, who we argue demonstrate enhanced top-down control over auditory processing (Kraus & Chandrasekaran, 2010; Strait et al., 2010). Although individuals who pursue music training may innately possess stronger efferent auditory control mechanisms than those who do not, we suggest that auditory advantages in musicians stem from focused and
consistent interactive experience with sound. We base this argument on correlations between years of music practice and auditory task performance (Kishon-Rabin et al., 2001; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Strait et al., 2010).

A Spotlight on Attention and Memory for Sound

Even in the quietest of rooms, our senses are inundated with a barrage of sound. From air ventilation systems to the hum of traffic, the human auditory system must adapt to a variety of listening conditions and hone in on signals of interest. What we hear is determined by how well we listen and by our capacity to retrieve what we’ve just heard from working memory, directing our attention to the input of highest interest while monitoring our surroundings for changes that require immediate attention. Focusing on a single auditory stream of interest becomes particularly challenging in the context of increasing noise levels, such as in the presence of multiple ongoing conversations (a cocktail party) or amidst environmental noise (a busy street). The adaptation of the human auditory system to a variety of listening environments is both impressive and essential for everyday human communication and auditory function. Deficits in auditory working memory and attention have been associated with a wide range of functional impairments such as auditory processing disorders (Moore, Ferguson, Edmondson-Jones, Ratib, & Riley, 2010), specific language impairment (Montgomery, 2002; Stevens, Sanders, & Neville, 2006) and developmental dyslexia (Facoetti et al., 2003; Jeffries & Everatt, 2004). In Principles of Psychology, William James asserts that "an education which should improve [attention] would be the education par excellence" (1890, p. 424). Our work and the work of others indicate that music training provides a mechanism for that very education in the auditory domain, enhancing our ability to direct our attentional spotlight, to remember what was recently heard, and to separate a target sound stream from other auditory input – not just for music but for other auditory domains as well.

Findings from our laboratory have demonstrated enhanced auditory attention and working memory in musicians (Figures 1b and c; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark et al., 2011; Strait et al., 2010). Increased auditory working memory in musicians has also been observed by others (Chan, Ho, & Cheung, 1998). The extent to which music training engenders an enhancement of general versus domain-specific (i.e., auditory-specific) working memory mechanisms is of interest. Whereas some studies report memory enhancements among musicians only for auditory tasks (Ho et al., 2003), others have observed carryover effects of music training in the visual domain (Jakobson, Lewycky, Kilgour, & Stoesz, 2008; see also Schellenberg, 2009). Through testing adult musicians and nonmusicians on auditory and visual working memory tasks, we found that adult musicians demonstrate enhanced auditory but not visual working memory compared to nonmusicians (Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark et al., 2011). This finding could not be attributed to a non-verbal IQ advantage in musicians. We have reported similar auditory-specific enhancements in musicians with regard to attention in that musicians demonstrate faster reaction times over the course of a sustained auditory attention task than nonmusicians (Figure 1b). They perform no better than nonmusicians, however, on an analogous visual attention task (Strait et al., 2010). Although we have not demonstrated visual working memory or visual attention enhancements in musicians compared to nonmusicians, some degree of visual cognitive enhancement would not be surprising given the multisensory nature of music practice and performance. In fact, our work and the work of others has revealed audiovisual perceptual (Petrini et al., 2009) and neural processing enhancements (Musacchia et al., 2007; Musacchia, Strait, & Kraus, 2008) in musically trained adults. In considering the domain specificity of working memory and attention enhancements in musicians, it is possible that auditory and visual working memory enhancements are each a matter of degree. Although music training may strengthen both processes, auditory cognitive abilities might be strengthened to a greater extent. Moreover, audiovisual processing enhancements in musicians may indicate that music training strengthens the perception and neural encoding of auditory and visual information, particularly when they co-occur. Given that memory and attention are multifaceted cognitive constructs with numerous subdivisions (e.g., attention can be sustained, phasic, selective, divided, or focused), considerable work is needed to tease apart the nature of music training’s impact on both cognitive functions and to determine potential carryover effects in other sensory domains.

How musicians perform on tasks that place a premium on auditory attention motivates hypotheses concerning the impact of music training on brain mechanisms that underlie auditory attention and working memory. During such tasks (e.g., when subjects are instructed to listen for certain target tones or timbres), musicians demonstrate heightened recruitment of cortical areas associated with sustained auditory attention and working memory (Baumann et al., 2007; Baumann, Meyer, & Jäncke, 2008;
Gaab & Schlaug, 2003; Haslinger et al., 2005; Pallesen et al., 2010; Stewart et al., 2003), such as the superior parietal cortex, as well as more consistent activation of prefrontal control regions (Strait & Kraus, 2011). Indications that music training increases contributions of the superior parietal and prefrontal cortices to active auditory processing and their roles in sustaining auditory attention and working memory may support the hypothesis that music training tunes the brain’s auditory cognitive networks for cross-domain auditory processing.

Not surprisingly, cortical networks invoked during attention to music are similar to those that underlie the activation of attention in other auditory domains, such as language. In addition to primary auditory areas (the superior temporal gyrus), these sites include the frontoparietal attention and working memory networks, which comprise the intraparietal sulcus, supplementary and presupplementary motor areas, and the precentral gyrus (Janata, Tillmann, & Bharucha, 2002). These findings corroborate previous results suggesting that a combination of modality-specific (e.g., superior temporal gyrus) and general attention and working memory centers (e.g., frontoparietal cortex) contribute to sustained auditory attention (Petkov et al., 2004; Zatorre, Mondor, & Evans, 1999). If music training increases musicians’ neural capacity for directing and sustaining auditory attention, music may prove to be the holy grail of auditory training, providing an avenue for the prevention, habilitation, and remediation of a wide range of auditory processing deficits.

**Enhanced Processing of Speech in Noise in Musicians**

**Attention and working memory: Key ingredients.** Given the frequency with which the human auditory system must encode signals of interest in challenging listening environments, the ability of the nervous system to process a target signal and suppress competing noise is essential for everyday human communication and auditory function. As noted, auditory attention and auditory working memory contribute to the ability to focus one’s auditory spotlight. Our laboratory has demonstrated associations between auditory attention/working memory performance and the ability to perceive speech in background noise (Figure 2a; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark et al., 2011; Strait & Kraus, 2011; see also Heinrich, Schneider, & Craik, 2008). Listening to speech in increasing levels of background noise requires augmented attentional resources, resulting in a decreased buffer capacity for auditory working memory storage and making quick recall of an entire sentence more difficult. Therefore, the greater one’s auditory attention and working memory, the better one’s ability to perceive speech in noise. Because of musicians’ enhanced auditory cognitive skills and their vast experience attending to and remembering distinct elements in complex soundscapes, we were not surprised to find that musicians demonstrate better perception of speech in background noise than nonmusicians. Specifically, musicians demonstrate an enhanced ability to repeat speech presented amidst louder levels of background noise (Figure 2a; Parbery-Clark, Skoe, Lam, & Kraus, 2009). That musicians’ speech-in-noise perceptual enhancement persists into the later decades of life is particularly relevant for older adults, who experience difficulty hearing speech in noise due to aging (Parbery-Clark et al., 2011; Zendel & Alain, 2011).

**Neurobiological evidence.** Neurobiological underpinnings of musicians’ enhanced perception of speech in noise have been documented in subcortical auditory processing. The human auditory brainstem response (ABR) provides a window into how complex sounds such as speech or music are transcribed into a neural code by the brain. Because ABRs preserve temporal and spectral components of evoking stimuli with exquisite temporal precision (on the order of fractions of a millisecond, as shown in Figure 3; see color plate section), they have proven useful for quantifying impairments and enhancements in human auditory processing, including enhancements in musicians. Because auditory brainstem neurons can phase lock to the frequency components of incoming sound up to ~2,000 Hz, auditory brainstem responses physically resemble the acoustic waveforms of evoking stimuli (see Figure 2b; Chandrasekaran & Kraus, 2010a). Subcortical encoding of the temporal features of incoming sound can be assessed by measuring the timing (i.e., latencies and phase) of neural response peaks that correspond to a sound’s onset, offset, rapid spectrotemporal changes, or sudden changes in level that occur with mid-stimulus amplitude bursts. The neural encoding of spectral components of incoming sound, such as its fundamental frequency and harmonics, and how these spectral components change over time can be assessed by measuring their representation within the ABR itself (via Fourier analysis, autocorrelation, or cross-correlations between the time-frequency components of the stimulus versus those in the ABR; Skoe & Kraus, 2010; Skoe, Nicol, & Kraus, 2011). Compared to quiet, ABRs to speech in background noise are slower and less robust (Burkard & Sims, 2002; Russo, Nicol, Musacchia, & Kraus, 2004). Notably, the amount of noise-induced degradation observed in ABRs is associated with behavior, with good speech-in-noise perceivers demonstrating less of a timing delay and more robust spectral encoding in noise (Anderson,
Compared to nonmusicians, musicians demonstrate decreased neural degradation to a speech stimulus (/da/) presented in background noise. Less neural degradation is seen in musicians' responses in both temporal and spectral dimensions (Figure 2c). Related results have been reported by Bidelman and Krishnan (2010) in

Chandrasekaran, Skoe, & Kraus, 2010; Parbery-Clark, Skoe, & Kraus, 2009; Song, Skoe, Banai, & Kraus, 2010). Compared to nonmusicians, musicians demonstrate decreased neural degradation to a speech stimulus (/da/) presented in background noise. Less neural degradation is seen in musicians' responses in both temporal and spectral dimensions (Figure 2c). Related results have been reported by Bidelman and Krishnan (2010) in
response to reverberated speech, with reverberation similarly complicating accurate speech perception and degrading speech-evoked ABRs. Musicians’ ABRs are more resistant to the degradative effects of stimulus reverberation on the subcortical encoding of speech, which points to their heightened ability to overcome challenging listening environments. That the degree of musicians’ speech-in–noise enhancement correlates with their duration of music training suggests that music training improves the ability of sensory systems to encode sound in challenging listening environments (Parbery-Clark, Skoe, Lam, & Kraus, 2009). A wealth of neurobiological studies have used correlational analyses to conclude that functional differences between the brains of musicians and nonmusicians are a consequence, at least in part, of music practice. In addition to predicting the perception of speech in noise, duration of training and its age of onset predict subcortical responses to speech, music, and emotional communication sounds (Musacchia et al., 2007; Strait et al., 2009; Wong et al., 2007), cortical structure (Gaser & Schlaug, 2003; Hutchinson, Lee, Gaab, & Schlaug, 2003) and cortical function (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Ohnishi et al., 2001; Pantev et al., 1998; Strait & Kraus, 2011; Trainor, Desjardins, & Rock, 1999). Although causality cannot be inferred from correlation, cortical (Margulis, Mlsna, Uppunda, Parrish, & Wong, 2009; Pantev, Roberts, Schulz, Engelen, & Ross, 2001; Shahin, Bosnyak, Trainor, & Roberts, 2003; Shahin, Roberts, & Trainor, 2004) and subcortical enhancements that are specific to the sound of one’s instrument of practice (Strait, Chan, Ashley, & Kraus, 2011) implicate the power of music training to shape brain function. Because experience-related and innate factors of musicianship undoubtedly coexist, future research needs to delineate their respective roles in shaping brain function. Longitudinal studies that assess different aspects of brain function before and after music training have been useful in this regard (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Hyde et al., 2009; Moreno et al., 2009; Norton et al., 2005; Schlaug, Forgeard, Zhu, Norton, & Winner, 2009) and are likely to provide the most informative outcomes.

Speech, Music, Rhythm and the Brain

Attentional involvement with sound can be mediated by sequential cueing, or rhythm (Jones, Kidd, & Wetzel, 1981). Rhythmic regularity provides listeners with a predictive framework, pushing their attentional spotlight to the most likely point in time where the next stimulus is likely to occur. That human nervous system tracks the occurrence of rhythmic signals (Snyder & Large, 2005), even when subjects are instructed to ignore them (Elhilali, Xiang, Shamma, & Simon, 2009), is not surprising because the human brain functions in inherently rhythmic ways. Rhythm, or periodicity, in neural activity is observed in electrophysiological recordings in the form of neural oscillations that are thought to drive a variety of cognitive functions and the synthesis of multisensory input (for review see Ward, 2003). Speech and music are likewise organized in similarly oscillatory manners. We now turn to a discussion of the importance of rhythm for the perception of both music and language. The available findings suggest that musicians’ experience producing, manipulating, and attending to musical sound promotes the brain’s ability to track regularities in sequential signals, even outside the domain of music (Conway, Pisoni, & Kronenberger, 2009).

Rhythm is a structural hallmark of both music and language. In music, rhythm has the power to drive our perception of phrase structure and tonality (Boltz, 1989) as well as to guide our anticipation of structural boundaries and cadences. Although less periodic than music, language is also structured metrically, with rhythm guiding the perception of speech prosody (e.g., syllabic stress; Marie et al., 2010) at rates that are consistent across all languages studied to date (Greenberg, 2003). Changes in speech at the syllable rate are fundamental to accurate speech perception (Drullman, Festen, & Plomp, 1994), with sensitivity to syllabic stress being a prerequisite for normal speech production and perception, as well as for the development of language-related skills. Indeed, the accurate perception of metrical structure in language is crucial to phonological development and, consequently, to the development of language-related skills such as reading (Corriveau, Pasquini, & Goswami, 2007; Goswami et al., 2002).

Cross-cultural study of the rhythmic structures of language and music suggests that the rhythm of a culture’s music is reflected in the rhythm of native-language prosody (Iversen, Patel, & Ohgushi, 2008; Patel, Iversen, & Rosenberg, 2006), implying that the two domains are inherently connected. Because music practice requires sustained attention to rhythm during the production and manipulation of instrumental output, cross-domain effects of music training on language perception would not be surprising. Enhanced sensitivity to the rhythmic components of speech could, in turn, promote syllabic discrimination and prosodic perception—two fundamental features of language processing that, when deficient, are associated with language and literacy dysfunction (Abrams, Nicol, Zecker, & Kraus, 2009; Goswami et al., 2002; cf. Tallal & Gaab, 2006).
Recent evidence points to an association between music perception—more specifically, rhythmic discrimination—and perceptual sensitivity to the metrical components of speech, such as syllabic stress (Huss, Verney, Fosker, Mead, & Goswami, 2010). This association points to the possibility of common mechanisms for processing sound in domains that depend on the sequential parsing of incoming information, enhancing that which is consistent (i.e., the metrical regularities of a target signal) to facilitate perception and stream segregation. In fact, sensitivity to regularities in sensory input is crucial for accurate perception in challenging listening environments. In a restaurant, for example, one must be sensitive to the regularities in an individual speaker’s voice, such as their rhythmic cadence, in order to focus on it amidst other simultaneously occurring conversations.

It is thought that the brain shapes perception according to higher-level predictions that it makes based on sensory regularities, sharpening sensory encoding at sequentially lower levels of the auditory pathway in a top-down manner (Ahissar, Lubin, Putter-Katz, & Banai, 2006; Ahissar, Nahum, Nelken, & Hochstein, 2009; Nahum, Nelken, & Ahissar, 2008). Recent work from our laboratory shows neural enhancements to regularities in an ongoing speech stream. Moreover, the extent of this enhancement relates to better performance on both music and language-related tasks, such as hearing speech in noise, reading, and music aptitude (more specifically, rhythm aptitude; Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Strait, Hornickel, & Kraus, 2011). We propose that, in combination with auditory cognitive abilities (working memory and attention), neural sensitivity to rhythmic components and regularities in ongoing sound streams provides a common mechanism that underlies music and reading abilities, potentially contributing to the observed covariance in child music and literacy skills (Anvari, Trainor, Woodside, & Levy, 2002; Atterbury, 1985; Forgeard, Schlaug, Norton, Rosam, & Iyengar, 2008; Overy, 2003; Strait, Hornickel, et al., 2011).

On the Association Between Music Training and Child Literacy

In our view, the auditory expertise gained over years of music training fine-tunes the auditory system in a comprehensive fashion, strengthening the neurobiological and cognitive underpinnings of both speech and music processing. Because of this, music training may promote the sensory and cognitive mechanisms that underlie child literacy (Besson, Schön Moreno, Santos, & Magne, 2007; Chandrasekaran & Kraus, 2010b, in press; Gaab et al., 2005; Huss et al., 2010). In fact, a growing body of work reports covariance in music and reading abilities (Forgeard et al., 2008; Overy, 2003), even after controlling for nonverbal IQ and phonological awareness (Anvari et al., 2002). Rhythm aptitude may be a better predictor of reading ability than pitch-based aptitude (Douglas & Willatts, 1994; Huss et al., 2010; Strait, Hornickel, et al., 2011), although earlier studies have been inconclusive in this regard (see Huss et al., 2010, for discussion; Forgeard et al., 2008; Overy, Nicolson, Fawcett, & Clarke, 2003; cf. Anvari et al., 2002).

In contrast to Hyde and Peretz’s (2004) classification of amusic brains as “out of tune but in time,” Goswami and colleagues have proposed that children with reading impairment have brains that are “in tune but out of time.” The connection between rhythm and reading abilities may reflect the fundamental importance of rhythm for both music and language production and perception. Indeed, we now have evidence for selective impairments in rhythm processing among individuals with dyslexia and Specific Language Impairment (SLI), including deficits in keeping a beat (Corriveau et al., 2007; Corriveau & Goswami, 2009; Goswami et al., 2002; Wolff, Michel, & Ovrut, 1990). Deficits in processing the fast temporal components of slow rhythmic events (i.e., syllables) have long been linked with reading impairment (Goswami, Fosker, Huss, Mead, & Szucs, 2011; Tallal, 1980; Tallal, Miller, & Fitch, 1993; Tallal & Stark, 1981) and the neural processing of rapid, non-linguistic sound sequences appears to be impaired in poor readers and children with SLI (Benasich & Tallal, 2002; Temple et al., 2000). Wisbey (1980) was one of the first to propose that music training, by facilitating the development of multisensory awareness and auditory acuity, could facilitate reading in impaired children. Although this proposal has been verified by a number of experiments (Douglas & Willatts, 1994; Moreno et al., 2009), more research using pre and post training paradigms is necessary to determine the specific impact of music training on child reading skill and underlying brain function, as well as to unravel contributions of tonal versus rhythmic aspects of music training.

Core neural and cognitive mechanisms that are deficient in poor readers are enhanced in musicians (Figure 3 in color plate section; Chandrasekaran & Kraus, 2010b; Kraus & Chandrasekaran, 2010; Kraus, Skoe, Parbery-Clark, & Ashley, 2009; Tzounopoulos & Kraus, 2009). Such mechanisms include auditory cognitive skills (working memory and attention; Figure 1b, c; Gathercole, Alloway, Willis, & Adams, 2006; Gathercole & Baddeley, 1990; Siegel, 1994), neural timing in response to speech (Figure 3 in color plate section, upper panel; Abrams et al., 2009; Banai et al., 2009; Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009; Musacchia et al., 2007, 2008; Parbery-Clark, Skoe, & Kraus, 2009; Strait et al., 2009), the neural
encoding of the spectral components of speech (Figure 3 in color plate section, middle panel), the neural discrimination of closely related speech syllables (Figure 3 in color plate section, lower panel; Chobert et al., 2011; Hornickel et al., 2009; Tierney, Parbery-Clark, Strait, & Kraus, 2011), and the perception and neural encoding of speech in noise (Figure 2; Anderson et al., 2010; Banai et al., 2009; Parbery-Clark, Skoe, & Kraus, 2009; Parbery-Clark, Skoe, Lam, & Kraus, 2009).

Because reading requires rapid and automatic mapping of written phonemes onto words stored in auditory memory, it is not surprising that how well the nervous system encodes speech sounds relates to reading ability (Banai et al., 2009; Chandrasekaran et al., 2009; Hornickel et al., 2009). Furthermore, increasing evidence indicates that some children with reading impairment are particularly susceptible to neural response degradation to speech in noise. For example, Anderson at al. (2010) demonstrated that, compared to good readers, children with reading impairment exhibit increased background noise-induced degradation in their neural responses to speech. By contrast, musicians demonstrate a relatively small impact of background noise on speech-evoked neural responses (Figure 2c; Parbery-Clark, Skoe, & Kraus, 2009). In light of connections between child reading impairment, the inability of the auditory system to encode a target signal of interest in the presence of noise (Sperling, Lu, Manis, & Seidenberg, 2005, 2006), impaired auditory cognitive skills (Gathercole et al., 2006; Gathercole & Baddeley, 1990; Siegel, 1994), and deficits in processing the fast temporal components of slow rhythmic events (i.e., syllables; Goswami et al., 2011; Tallal, 1980; Tallal et al., 1993; Tallal & Stark, 1981), we suggest that music may provide an efficient mechanism for auditory training, aiding in the prevention and remediation of speech-in-noise and reading impairments (Chandrasekaran & Kraus, 2010b; Kraus & Chandrasekaran, 2010).

Conclusions

Due to its multisensory nature, attentional demands, complex sound structure, rhythmic organization, and reliance on rapid audio-motor feedback, music is a powerful tool for shaping neuronal structure and function, especially with regard to auditory processing (Norton et al., 2005; Schlaug, 2001; Schlaug et al., 2009; Schlaug, Norton, Overy, & Winner, 2005). Its effects are not constrained to the brain's music networks but apply to general auditory processing, including the processing of speech. We have herein defined common mechanisms that underlie music and language skills such as reading and hearing speech in noise. These mechanisms include auditory attention, working memory, neural function in challenging listening environments (e.g., in background noise), sequential sound processing, and neural sensitivity to temporal and spectral aspects of complex sounds and sound regularities. Given these common mechanisms, music may be useful for promoting the development and maintenance of auditory skills and for improving the efficacy of remedial attempts for individuals with auditory impairments.

Author Note

This research is supported by National Science Foundation 0921275 and 1057556, the National Institutes for Health F31DC011457-01 and by the Grammy Research Foundation.

Correspondence concerning this article should be sent to Nina Kraus, Auditory Neuroscience Laboratory (www.brainvolts.northwestern.edu), Northwestern University, 2240 Campus Drive, Evanston, IL 60208. E-mail: nkraus@northwestern.edu

References


A cortical “language-network” serves the processing of music.  
NeuroImage, 17, 956–966.


Training-induced malleability in neural encoding of pitch, timbre and timing: Implications for language and music.  

The effects of tone language experience on pitch processing in the brainstem.  
Journal of Neurolinguistics, 23, 81–95.

Encoding of pitch in the human brainstem is sensitive to language experience.  

Selective subcortical enhancement of musical intervals in musicians.  
Journal of Neuroscience, 29, 5832–5840.

Selective neurophysiologic responses to music in instrumentalists with different listening biographies.  


Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiologcal evidence.  
Journal of Cognitive Neuroscience, 19, 1453–1463.

Influence of musical and psychoacoustical training on pitch discrimination.  
Hearing Research, 219, 36–47.

MONTGOMERY, J. W. (2002). Working memory and comprehension in children with specific language impairment:  
What we know so far.  
Journal of Communication Disorders, 35, 221–231.

Nature of auditory processing disorder in children.  
Pediatrics, 126, e382–390.

Cerebral Cortex, 19, 712–723.

Musicians have enhanced subcortical auditory and audiovisual processing of speech and music.  

Hearing Research, 241, 34–42.

PloS Biology, 6, e126.

Are there pre-existing neural, cognitive, or motoric markers for musical ability?  

Functional anatomy of musical perception in musicians.  
Cerebral Cortex, 11, 754–760.


Dyslexia and music: Measuring musical timing skills.  
Dyslexia, 9, 18–36.

Cognitive control in auditory working memory is enhanced in musicians.  
PloS ONE, 5, e11120.


Timbre-specific enhancement of auditory cortical representations in musicians.  
Neuroreport, 12, 169–174.

Musical experience limits the degradative effects of background noise on the neural processing of sound.  
Journal of Neuroscience, 29, 14100–14107.

Musician enhancement for speech-in-noise.  
Ear and Hearing, 30, 653–661.

Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise.  
PloS ONE, 6, e18082.

Nature Neuroscience, 6, 674–681.

Comparing the rhythm and melody of speech and music: The case of British English and French.  

Stronger bilateral efferent influences on cochlear biomechanical activity in musicians than in non-musicians.  

Attentional modulation of human auditory cortex.  


