



Neurophysiological influence of musical training on speech perception

Antoine J. Shahin*

Department of Otolaryngology-Head and Neck Surgery, The Ohio State University, Columbus, OH, USA

Edited by:

Lutz Jäncke, University of Zurich, Switzerland

Reviewed by:

Erich Schröger, University of Leipzig, Germany

Nina Kraus, Northwestern University, USA

***Correspondence:**

Antoine J. Shahin, Department of Otolaryngology-Head and Neck Surgery, The Ohio State University, 915 Olentangy River Road, Columbus, OH 43212, USA.
e-mail: shahin.5@osu.edu

Does musical training affect our perception of speech? For example, does learning to play a musical instrument modify the neural circuitry for auditory processing in a way that improves one's ability to perceive speech more clearly in noisy environments? If so, can speech perception in individuals with hearing loss (HL), who struggle in noisy situations, benefit from musical training? While music and speech exhibit some specialization in neural processing, there is evidence suggesting that skills acquired through musical training for specific acoustical processes may transfer to, and thereby improve, speech perception. The neurophysiological mechanisms underlying the influence of musical training on speech processing and the extent of this influence remains a rich area to be explored. A prerequisite for such transfer is the facilitation of greater neurophysiological overlap between speech and music processing following musical training. This review first establishes a neurophysiological link between musical training and speech perception, and subsequently provides further hypotheses on the neurophysiological implications of musical training on speech perception in adverse acoustical environments and in individuals with HL.

Keywords: auditory cortex, speech perception, musical training, hearing loss, speech in noise, EEG, MEG, neuroplasticity

The brain is astoundingly neuroplastic, and studying neurophysiological alterations due to musical training is an increasingly popular method to assess the precise mechanisms underlying neuroplasticity. Furthermore, it is particularly interesting to understand how neuroplastic changes due to musical training may affect other auditory functions, namely speech perception. It has been repeatedly shown that musical training modifies the auditory neural circuitry in a way that allows for enhanced musical skills (e.g., processing musical timbre, pitch contour, and rhythm). However, the extent to which enhanced auditory representations and processing due to musical training transfers to other auditory functions, especially speech perception, is an issue of exciting debate and a rich topic to explore. In order for enhanced auditory skills due to musical training to transfer to speech perception, musical training must target the neural mechanisms that underlie speech perception and modify them in a way to enhance their function.

From a psychoacoustics perspective, music and speech processing share several mechanisms. To extract meaning from a piece of music or ongoing speech, the auditory system must encode pitch or voice cues (to identify a musical note or speaker) as well as rhythm (the unfolding of words and chords over time in a discourse or symphony), and segregate a voice in the midst of background noise or a musical instrument within a symphony orchestra. Thus, the neural mechanisms involved in processing music should overlap with and influence speech mechanisms (and vice versa). Furthermore, musical training should recruit resources associated with speech processing, thereby enhancing the neural mechanisms underlying speech perception (and vice versa). This article reviews neurophysiological evidence supporting an influence of musical training on speech perception at the sensory level, and discusses whether such transfer could facilitate speech perception in individuals with hearing loss (HL).

METHODOLOGIES FOR ASSESSING AUDITORY STRUCTURE AND FUNCTION

Effects of musical training on brain attributes have been largely studied using non-invasive imaging methods, primarily electroencephalography (EEG), magnetoencephalography (MEG), magnetic resonance imaging (MRI), and functional MRI (fMRI). However, the electrophysiological approaches (EEG and MEG) will be emphasized in this review.

Electroencephalography and MEG detect the brain's electrical and magnetic activity, respectively, at the scalp with a millisecond resolution. The amplitude of the event-related potential (ERP as in EEG) or field (ERF as in MEG) or of the auditory-evoked potential/field (AEP/AEF) indicates the size of the activated neural population and/or the trial-to-trial phase synchrony of neural firing when evoked by sensory stimuli (e.g., sound). The latency of the ERP/ERF reflects the timing of neural activation. Combined with the topography of the neural response (i.e., its position on the scalp), these ERP/ERF signatures provide specifics on the function and origin of activity (Pantev et al., 2001; Musacchia et al., 2008). MRI and fMRI, in contrast, probe the entire brain as opposed to scalp activity. MRI provides fine details of brain structure due to differences in magnetic properties of brain tissues (i.e., white matter versus gray matter). fMRI reveals the hemodynamic response associated with neural activity, and the foci of this activity are then superimposed on MRI images to determine the precise neuroanatomical structure(s) generating the neural activity. Thus, both MRI and fMRI reveal spatially well-defined neuroanatomical and functional loci undergoing structural and functional changes, for example due to musical training (Ohnishi et al., 2001; Hyde et al., 2009).

PREREQUISITES FOR FUNCTIONAL TRANSFER FOLLOWING ACOUSTICAL TRAINING

Acoustical features of speech and music are represented by a hierarchical auditory network. Low-level regions [i.e., brain stem and primary auditory cortex (A1)] encode simple acoustical features, such as sound onset and pitch. More complex features, such as spectrotemporal combinations that represent speech or musical timbre, recruit higher-level processes in non-primary auditory cortex (NPAC). Studies have provided evidence for neural networks, at the level of the auditory cortex and beyond, that are relatively specialized for processing either music or speech (Peretz et al., 1994; Tervaniemi et al., 1999, 2006; Zatorre et al., 2002; Rogalsky et al., 2011). However, functional and structural overlap for speech and music processing along the sensory (Wong et al., 2007; Sammler et al., 2009; Rogalsky et al., 2011) and cognitive (Sammler et al., 2009; Schulze et al., 2011) levels exist. A prerequisite for the transfer of auditory function to speech perception following musical training is the fostering of greater overlap between the neural mechanisms associated with speech and music perception. To put this into perspective consider the following example: since a sound's temporal information, which is fundamental to speech perception, is favorably processed in the left auditory cortex (Zatorre and Belin, 2001; Zatorre et al., 2002), musicians may be inclined to regularly use left-hemisphere resources to assess temporal relationships between musical segments (Bever and Chiarello, 2009). This in turn may induce neuroplastic modifications in the left auditory cortex that would support greater temporal processing for speech stimuli in musicians than in non-musicians.

From an electrophysiological perspective (EEG/MEG), musical training or training in the speech domain, such as learning to discriminate slight differences in the fundamental frequency (f_0) of vowels (Reinke et al., 2003) or voice onset time (VOT; Tremblay et al., 2001), has been linked to neuroplastic modifications of the same neural components associated with processing acoustical features common to music and speech. Thus, both forms of training may lead to an increase in overlap of the neural mechanisms that underlie speech and music processing, and thus a transfer of auditory function between the two domains may be possible. A caveat of this assumption is that even if training in speech or music leads to modification of the same neural component(s) measured in EEG/MEG, one cannot necessarily assume that the observed neuroplastic changes occurred in overlapping neural populations. The use of fMRI or source localization techniques in EEG/MEG (Freeman and Nicholson, 1975; Scherg, 1990; Hamalainen and Ilmoniemi, 1994; Tervaniemi et al., 1999) would elucidate the extent of such neural overlap. However, if training in one domain (e.g., music) influences the same neural processes in the other domain (e.g., speech), emerging as a neuroplastic change in the same EEG/MEG components, then an overlap in processing and thus a transfer of function can be inferred. In this review, I will first present evidence demonstrating that musical training and training in the speech domain may target the same neurophysiological components along the auditory pathway. Next, I will show evidence of musical training-related modulations of these same components during speech processing (i.e., transfer of auditory function). Finally, I will discuss how musical training might enhance speech perception in adverse acoustical environments and in individuals with HL.

MUSICAL TRAINING AND TRAINING IN THE SPEECH DOMAIN TARGET SHARED NEURAL MECHANISMS

BRAIN STEM

Training in both musical and speech domains affect brain stem processing, as evidenced by training-related changes in the EEG frequency following response (FFR). The FFR reflects the resonance (amplitude and phase-alignment) of brain stem neuronal firing with a sound's f_0 . Therefore, a more robust FFR following training in music and speech is indicative of an enhanced representation of f_0 at the brain stem (Musacchia et al., 2007; Krishnan et al., 2009). Indeed, the FFR is more robust in musicians for music sounds than in non-musicians (Musacchia et al., 2007). Likewise, in the realm of speech learning, the FFR is enhanced in English speakers for Mandarin tones, following training on tonal speech (Song et al., 2008). Taken together, tonal training, either in music or speech, elicits neuroplastic modifications in the brain stem FFR, which implies that an inter-domain functional transfer at the brain stem level is probable.

PRIMARY AUDITORY CORTEX (A1)

While training in both speech and musical domains target the FFR, there is less evidence for such functional overlap at the level of A1, as indexed by the middle latency response (MLR). The MLR, which is measured in EEG and MEG, is believed to reflect thalamo-cortical input to and processing in A1 (Hall, 2006). The MLR consists of Na/Pa/Nb/P1 components (see **Figure 1**; Note that the P1 is also referred to as the P50) and typically occurs between 19 and 50 ms following sound onset. Larger MLR components were shown to index enhanced pitch and rhythm encoding in musicians relative to non-musicians for pure tones and music sounds (N19m–P30m – magnetic counterparts of the electrical Na–Pa: Schneider et al., 2002; P1: Shahin et al., 2004; Neuhaus and Knösche, 2008). Furthermore, a training-related increase in P1 amplitude may reflect the acoustical feature binding mechanisms that integrate rhythmic and pitch patterns into

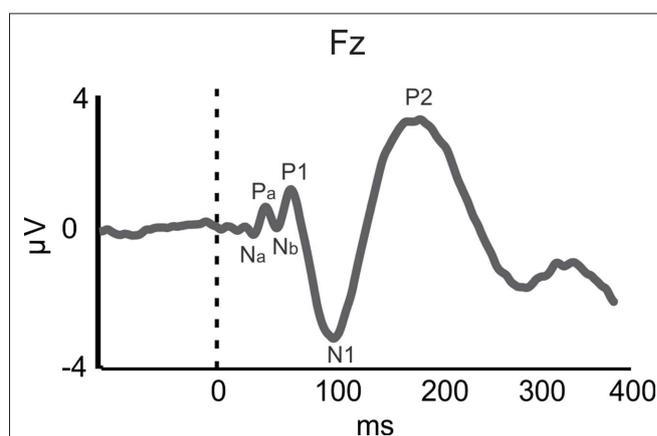


FIGURE 1 | Auditory-evoked potentials (AEPs). AEP waveform for piano tone at the frontal channel Fz for 34 participants (averaged over musicians and non-musicians). The Na–Pa–Nb–P1 complex represents the middle latency response (MLR) originating in primary auditory cortex (A1). The N1 and P2 components represent activity originating in the surrounding belt areas of A1. Based on data from Shahin et al. (2003).

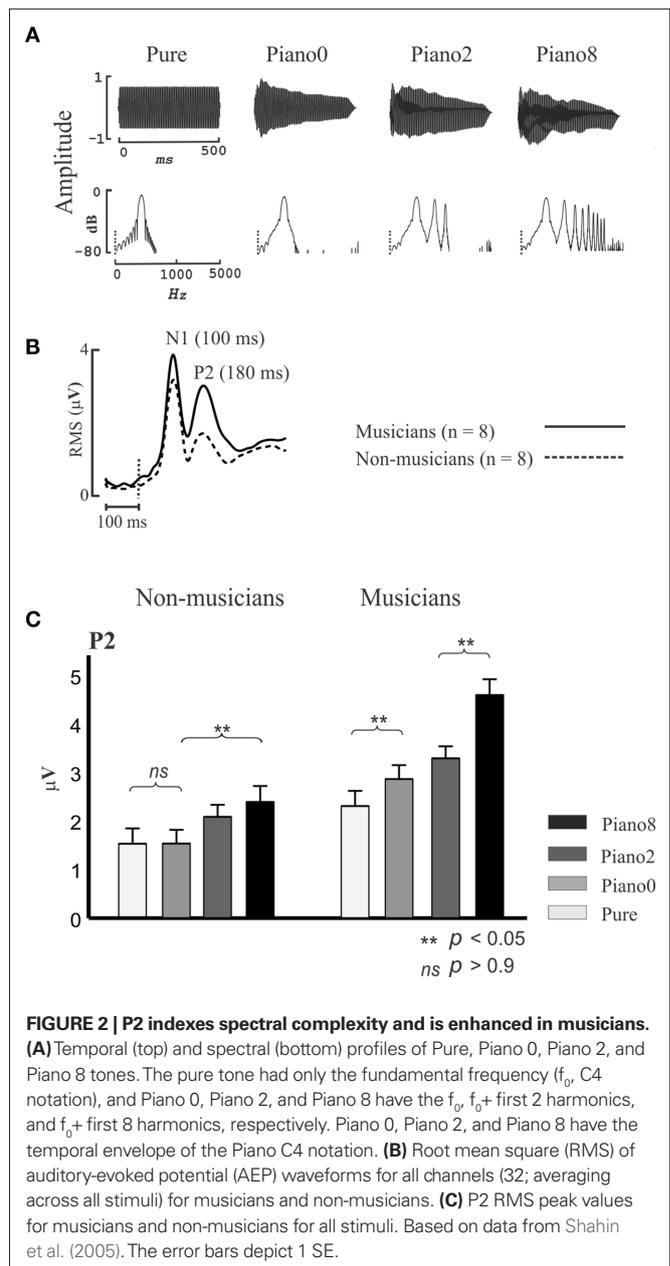
a coherent melody (Neuhaus and Knösche, 2008). However, a recent study challenged the relation between the P1 and musical expertise, such that the P1 was smaller in amplitude and delayed in musicians than in non-musicians for harmonic tones (Nikjeh et al., 2009).

In the realm of speech perception and training in the speech domain, the functional role of the MLR is less clear. Since we are more exposed to voiced than unvoiced speech in naturalistic situations and musical training affects the MLR, one would expect that voiced stimuli would elicit larger MLR amplitude than unvoiced stimuli, which lack pitch information. However, Hertrich et al. (2000) showed that the M50 (the magnetic counterpart of the electrical P1) was reduced for periodic (voiced) compared to aperiodic (unvoiced) speech-like stimuli. Instead, the M100 (the magnetic counterpart of the electrical N1), which has sources in NPAC, was larger for periodic than non-periodic speech-like stimuli (Hertrich et al., 2000). This suggests that experience-based enhancements in pitch processing of speech may commence at a higher cortical level than that of music. Likewise, discrimination training on VOT in young adults resulted in smaller P1 amplitudes but larger N1 and P2 AEPs (Tremblay and Kraus, 2002). Taken together, these results imply that higher cortical regions (i.e., NPAC) favor the processing of more spectrotemporally complex signals, such as speech, consistent with animal (Rauschecker et al., 1995), and human fMRI (Rauschecker et al., 1995; Patterson et al., 2002) results. Thus, whereas enhanced MLR may reflect neuroplastic modifications in temporal (i.e., rhythm) and spectral (i.e., pitch) processing following musical training, analogous MLR effects in the speech domain do not seem to be related to expertise. Instead, expertise-related effects in the speech domain appear to emerge later, in the N1 (~100 ms) and P2 (~180 ms; **Figure 1**) AEPs/AEFs.

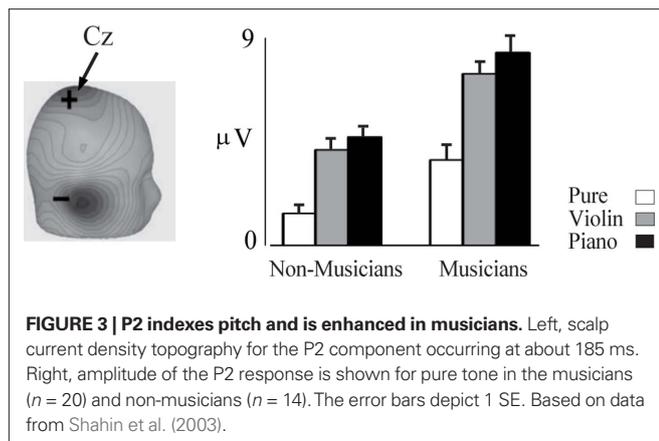
NON-PRIMARY AUDITORY CORTEX

Because of the complex spectrotemporal structure of music and speech sounds, their processing may be favored in NPAC. Thus, the possibility of inter-domain functional transfer is more likely to occur in later auditory processing stages (i.e., N1 and P2) instead of the MLR.

The N1 and P2 are thought to originate in the region surrounding A1 (Shahin et al., 2003; Bosnyak et al., 2004), including belt and parabelt regions of the superior temporal gyrus (Hackett et al., 2001). These regions are collectively referred to here as NPAC. While the N1 and P2 code for low-level acoustical features such as sound onset and pitch, they also represent higher-level sound features brought about by the spectrotemporal complexity of speech and music. **Figure 2** shows the P2 AEP as a function of sound complexity. Musicians and non-musicians were presented with four tones varying in spectral complexity. Three of the tones were piano tones (C4 note) which contained the fundamental (f_0) and 8 (Piano 8), 2 (Piano 2), or no (Piano 0) harmonics, and a pure tone (Pure) which had the same f_0 but not envelope of the piano tone. Notice that as more harmonics were added to the piano tone (**Figure 2A**), P2 increased in amplitude [**Figure 2C**; 2 (group) \times 4 (tone) ANOVA; group main effect $F_{(1,14)} = 14, p < 0.005$; tone main effect $F_{(3,42)} = 35, p < 0.0001$, interaction $F_{(3,42)} = 6, p < 0.001$], especially in musicians [*post hoc* Fisher's least significant difference (LSD) test $p < 0.05$]. Also, the P2 amplitude in response to the Piano 0 tone



was significantly larger than the P2 for the pure tone in musicians ($p < 0.02$, LSD test) but this difference was not evident in non-musicians. Pure and Piano 0 tones differed only in the shape of their temporal envelopes, with Piano 0's envelope matching the temporal envelope of the piano C4 note, thus implicating the P2 in coding the temporal onset of the learned sound. Finally, although the P2 amplitude for the pure tone tended to be larger in musicians than in non-musicians in this study ($p < 0.07$ LSD test), the same contrast was highly significant when professional musicians were used in a prior study (**Figure 3**). P2 enhancement for pure tones in musicians represents enhanced pitch encoding, even in a non-timbre context, and hence may index basic transfer of auditory function due to musical training. In sum, the P2 is associated with coding the temporal (temporal onset) and spectral (harmonics) features of



music sounds, and these functions are refined by musical training. Like the P2, the N1 was also associated with increased number of harmonics and with coding the timbre in ERP/ERF studies (Pantev et al., 2001; Seither-Preisler et al., 2003; Meyer et al., 2006).

More pertinent to the current topic is the finding that acoustical training in the speech domain has been shown to evoke larger P2s for the same acoustical features that led to P2 augmentation following musical training. Short-term training on vowel discrimination tasks based on slight f_0 manipulations elicited a larger P2 responses in the trained group compared to the untrained group (Reinke et al., 2003), in the same way that pure tones matched in f_0 to music tones evoked a larger P2s in musicians than in non-musicians (Shahin et al., 2003). Also, P2 enhancement was reported following VOT discrimination training (Tremblay et al., 2001, 2009; Tremblay and Kraus, 2002) in the same manner that musicians exhibited larger P2s to the temporal onset of learned music sounds (Figure 2). These increases in P2 amplitude observed following either musical training or training in the speech domain may index a greater overlap of neural populations, via increased neuronal recruitment associated with processing spectral or temporal features of music and speech. In short, if musical training or training in the speech domain can foster neuroplastic changes observed in the same brain response (i.e., P2) associated with processing the same acoustical feature, enhanced speech processing due to musical training is likely.

It is worth noting that the N1's and P2's relationship to coding the spectrotemporal combination of acoustical features (timbre) may suggest a feature binding role for these AEP components. In support of this hypothesis, Neuhaus and Knösche (2008) showed the interdependency of the neuroplastic N1 and P2, as well as the earlier P1 component, on pitch and rhythm processing. They presented musicians and non-musicians with melodies, either ordered by pitch and duration, or randomized in pitch, duration, or both. They found a general effect in which the amplitudes of the P1, N1, and P2 AEPs were augmented as the level of disorder in pitch and/or duration increased in both musicians and non-musicians (Neuhaus and Knösche, 2008). Furthermore, in musicians but not in non-musicians, the P1 and P2 amplitudes were influenced (augmented) by pitch randomization only when the duration was ordered, suggesting that the integration of pitch and rhythm processing is fostered by musical training. Thus, the results imply that an increase

in these AEP amplitudes may reflect increased neural effort associated with binding spectrotemporal features into a coherent melody. Feature binding may be facilitated by selective attention (Treisman and Gelade, 1980) via modulation of early auditory cortical activity. However, increased spectrotemporal acuity observed in musicians (as discussed previously) must be a precursor of enhanced feature binding, whether automatic or under attentive conditions. In short, interdependence of pitch and rhythmic processing and thus binding of these acoustical features may be fostered by musical training at the level of A1 and NPAC (P1, N1, and P2 AEPs). A remaining question is whether enhancements in feature binding mechanisms due to musical training are specific to the learned auditory objects (Pantev et al., 2001) or are also transferable to speech perception.

NEUROPHYSIOLOGICAL EVIDENCE FOR THE INFLUENCE OF MUSICAL TRAINING ON SPEECH PERCEPTION

The previous section highlighted neurophysiological evidence supporting targeted modification of the same neural components (FFR and P2) by musical training and training in the speech domain. These findings do not provide evidence for inter-modal functional transfer, but rather offer a starting point in formulating hypotheses regarding the likelihood of such transfer. In this section, I will present findings showing that musical training can impact speech perception (i.e., inter-modal functional transfer), through neuroplastic changes in the same components that were modified by training in either the music or speech modality (previous section).

MUSICAL TRAINING'S INFLUENCE ON PITCH AND RHYTHM ENCODING IN SPEECH PERCEPTION

Superior pitch and rhythm encoding for speech in musicians, compared to non-musicians, is reflected at the brain stem and auditory cortex.

Brain stem

At the brain stem level, the FFR was shown to be more robust in musicians than in non-musicians for both music and speech (/da/) stimuli (Musacchia et al., 2007, 2008). Furthermore, enhanced pitch tracking (FFR) for Mandarin speech at the brain stem was seen in English-speaking (unfamiliar with tonal languages) musicians compared to non-musicians during passive listening to Mandarin words (Wong et al., 2007). The aforementioned neurophysiological evidence was consistent with better pitch discrimination abilities in musicians compared to non-musicians (Musacchia et al., 2007, 2008; Wong et al., 2007). More precise pitch encoding of speech sounds in musicians at the brain stem and under passive listening conditions (Wong et al., 2007) suggests that functional transfer of auditory mechanisms to speech perception following musical training can occur even at a relatively low level of the auditory pathway. However, corticofugal feedback from the cortex cannot be discounted, especially in attentive conditions (Musacchia et al., 2008).

Auditory cortex

The aforementioned enhanced FFR effect reported by Musacchia et al. (2008) for the speech syllable/da/was accompanied by earlier latencies and larger amplitudes of the components of the P1–N1–P2–N2 complex, which have origins in A1 and NPAC. The P2

enhancement in musicians compared to non-musicians is noteworthy. Musacchia et al.'s (2008) results are consistent with the neurophysiological consideration that if musical training and training in the speech domain impact a shared neural system similarly – as observed via P2 enhancement following musical (Pantev et al., 1998; Shahin et al., 2003) and vowel and VOT discrimination training (Tremblay and Kraus, 2002; Reinke et al., 2003) – then functional transfer to speech perception following musical training can be inferred. A more recent study showed a similar P2 effect that implicated rhythm processing as the mechanism affected by functional transfer following musical training. Specifically, Marie et al. (2010) demonstrated that the neuroplastic P2, previously shown to reflect temporal acuity (Tremblay et al., 2001; Figure 2), is associated with the musicians' greater ability to encode metric structure in speech. Marie et al. (2010) varied syllabic length of the last word in a collection of sentences and measured ERPs while musicians and non-musicians judged whether the last word in a sentence was well-pronounced. They found that compared to non-musicians, musicians exhibited larger P2 amplitudes, coupled with greater accuracy, for metrically incongruous than for congruous words (Marie et al., 2010). While this P2 enhancement is consistent with a transfer effect associated with coding the temporal acuity of speech, it may also be partly attributed to increased perceptual effort to integrate the sequential pattern into a coherent segment (Neuhaus and Knösche, 2008). In sum, these studies demonstrate that P2 enhancement, an index of musical training (Shahin et al., 2003), for coding the metric structure in speech coincides with enhanced temporal (Tremblay et al., 2001; Marie et al., 2010) acuity in speech processing. Thus, a transfer effect for spectrotemporal features, and possibly the neural mechanisms involved in binding them, can be facilitated by musical training.

In sum, there is preliminary neurophysiological evidence, exhibited in several obligatory auditory components (e.g., FFR, P2), supporting a functional transfer to speech perceptual mechanisms following musical training. Moving forward, it is important to weigh the implications of such neurophysiological transfer to speech perception in naturalistic acoustical environments and in individuals with HL.

MUSICAL TRAINING'S INFLUENCE ON SPEECH PERCEPTION IN ADVERSE ACOUSTICAL ENVIRONMENTS

One important aspect of speech processing related to everyday auditory experience, especially for populations with HL, is the ability to comprehend speech in acoustically adverse environments. For example, in a cocktail party, a listener must segregate her friend's voice from a concurrent multi-talker background to understand his story. Musicians are better at concurrent sound segregation (Zendel and Alain, 2009), which entails enhanced pitch discrimination ability (Tervaniemi et al., 2005), working memory function (Chan et al., 1998), and selective attention (Strait et al., 2010). Thus, musically trained individuals should demonstrate improved perception of speech in everyday noisy environments. Indeed, when contrasting speech comprehension abilities of musicians and non-musicians with comparable hearing status, musicians perform better in noisy backgrounds (Parbery-Clark et al., 2009b). Parbery-Clark et al. (2009b) administered the hearing-in-noise test (HINT), the quick speech-in-noise (QuickSIN) test, frequency discrimination

tests, and working memory tests to musicians and non-musicians. HINT tests speech perception in a background speech-shaped noise, while QuickSIN tests speech perceptual ability in a multi-talker environment. Musicians outperformed non-musicians in all tasks. Furthermore, musicians' performance on all tests, except for HINT, correlated with years of practice. These results suggest that enhanced working memory, selective attention, and frequency discrimination abilities as a function of musical expertise play a significant role when the target speech is embedded within a multi-talker environment (QuickSIN) versus when background noise is environmental (HINT).

The aforementioned behavioral outcomes seen in musicians (Parbery-Clark et al., 2009b) were subsequently supported by electrophysiological findings. Parbery-Clark et al. (2009a) showed that the brain stem response morphology was less influenced by background noise in musicians than in non-musicians. They examined the latency and amplitude of two auditory brain stem response (ABR) peaks representing the onset of the consonant–vowel pair/da/ and the formant transition (/d/to/a/) in musicians and non-musicians and in quiet and multi-talker babble (Parbery-Clark et al., 2009a). When comparing results of noisy and quiet conditions, the authors found that the latencies of the ABR peaks were less delayed and the amplitudes were more preserved in musicians than in non-musicians. The authors concluded that the degradative effects of noise on the neural processing of sounds can be limited with musical training.

It is not clear whether musical training can enhance other forms of auditory scene analysis. For example, does musical training boost one's ability to perceive degraded speech, as in phonemic restoration (PR)? In PR, a word with a missing segment can be perceived as continuous, provided that the missing segment is masked by another sound, such as a cough or white noise (Warren, 1970). The ability to restore speech in PR requires a dynamic mechanism in which the missing spectrotemporal structure is interpolated (also known as perceptual “filling-in”) from the masking entity. In addition, A1's usual sensitivity to the onsets/offsets of missing segments (Riecke et al., 2009; Shahin et al., 2009) is suppressed, resulting in the illusory perception of continuous speech. This filling-in process relies on higher-level mechanisms, such as Gestalt processing and access to prior knowledge (e.g., template matching in memory; Shahin et al., 2009). The question then becomes, how would musical training affect PR? Musicians may be expected to perform worse (i.e., fail to fill-in the degraded speech signal) than non-musicians on a PR task, given their enhanced concurrent sound segregation abilities (Zendel and Alain, 2009) and stronger sensitivity to acoustical onsets (Musacchia et al., 2008) and offsets of missing segments. However, since PR strongly relies on top-down influences, it can be argued that musical training may benefit PR.

Two analogies to PR include restoration of the missing fundamental (f_0) and of the missing beat in a metric structure (Bendor and Wang, 2005; Snyder and Large, 2005). Representations of an omitted f_0 in the auditory cortex (Bendor and Wang, 2005) are preserved due to interpolation from the remaining harmonic structure of the sound. Yet this skill (restoration of f_0) has been shown to be more robust in musicians than in non-musicians (Preisler, 1993; Seither-Preisler et al., 2007). However, restoration of a missing beat within a metric structure is more closely related to PR,

since like PR it involves filling-in in the temporal domain (Snyder and Large, 2005). Snyder and Large (2005) showed that when a tone was omitted from a rhythmic train of tones, auditory gamma band activity continued through the gap as if the missing tone was physically present, thus preserving the cortical representations of the rhythmic tone pattern. This temporal filling-in is likely related to enhanced Gestalt integration, expectancy, and template matching in long-term memory, which are all processes important in PR and are enhanced in musicians compared to non-musicians (Besson and Faita, 1995; Fujioka et al., 2005; Lenz et al., 2008; Shahin et al., 2008). An experiment assessing the sensitivity of A1 to the onsets and offsets of missing speech segments in musically trained and non-trained individuals should determine how musical training influences PR. Reduced sensitivity of A1 (i.e., reduced MLR amplitudes) in musicians than in non-musicians to onsets/offsets of missing segments would support the premise that musical training enhances PR, thereby providing convincing evidence of the transfer of auditory function to speech perception following musical training.

IMPACT OF MUSICAL TRAINING ON SPEECH PERCEPTION IN INDIVIDUALS WITH HEARING LOSS

As stated earlier, making sense of degraded speech in noisy environments is one of the most critical challenges for individuals with HL. Although individuals with HL (e.g., presbycusis) can typically understand speech in quiet settings (e.g., a one-on-one conversation with little or no background noise), degraded speech or the addition of background noise has a disastrous effect on their speech comprehension. The etiology of HL varies considerably among individuals. In general, HL in the aging population (presbycusis) is characterized by difficulty hearing high-frequency sounds, while individuals with cochlear implants (CIs) must make sense of CI signals that are severely limited in spectral details. If musical training modifies the neural circuitry of speech processing, then how would such modification affect speech perception experience in individuals with HL?

Loss of sensitivity to high-frequency sounds in the aging population leads to a reduced ability to perceive certain sounds, such as the phonemes *s*, *sh*, and *ch*. For example, assuming intact low-frequency perception, the sentence *Trisha cherishes her friend's memories* may sound as *Tri**a**eri**e* her friend** memorie**, where the asterisks indicate unintelligible sounds. Thus, similar to PR, individuals with HL must temporally fill-in representations of degraded speech. If musical training enhances PR mechanisms as discussed above, then musical training could serve as a possible prevention/intervention strategy in individuals with presbycusis. Neurophysiologically, this can be assessed by examining A1 sensitivity to onsets/offsets of degraded fricatives (e.g., *s*, *sh*) in musician and non-musician groups of older individuals with similar HL etiology.

In the case of individuals with CIs, the problem is related to how musical training can improve perception of the spectrally degraded signal outputted by the CI. CIs convert a speech signal into a few (usually eight or more) channels that represent several frequency bands of noise. The signals outputted by CIs can be imitated closely, though not perfectly, by noise-vocoded speech. In vocoded speech, the spectral structure of speech is replaced by noise bands, resulting in a highly degraded spectral structure (i.e., pitch quality), but the

overall amplitude (i.e., temporal envelope) of the sound is preserved to a good degree. It is not surprising that CI users rely on temporal information (rhythm) to a greater extent than a sound's spectral information (pitch; Gfeller and Lansing, 1991). Musical pitch perception in CI users has been shown to be highly correlated with lexical pitch perception, suggesting shared neural mechanisms for processing pitch in music and speech in CI users (Wang et al., 2011). Given that musical training refines one's ability to resolve fine pitch differences in normal-hearing individuals, it would be plausible that CI users who receive musical training may show improvements in pitch interval perception compared to non-musically trained CI individuals. Indeed, a recent study revealed that musically trained children with congenital/prelingual deafness fitted with CIs exhibit better pitch identification than non-trained children (Chen et al., 2011). Furthermore, pitch identification accuracy in the trained children was correlated with the duration of musical training. However, the musical training-related functional transfer to speech perception seen in normal-hearing individuals (discussed above) has not been demonstrated in CI users. A recent longitudinal study using behavioral measures found that while musically trained (over a 2-year period) CI children performed better on pitch discrimination tasks than children who did not receive music lessons, the two groups performed equally on speech discrimination tasks (Yucel et al., 2009). Additional studies using electrophysiological measures may be necessary, as changes due to musical training, especially over short time periods, may emerge more quickly in subcortical and cortical assessments, preceding observed behavioral improvements. This can be examined using pitch-sensitive auditory-evoked components (P2 AEP, Reinke et al., 2003) to assess pitch perception fidelity in speech and speech in noise following musical training in individuals with CIs. Longer-term musical training should also be considered in HL populations.

Another practical approach to explore involves assessing the effects of musical training on audio-visual (AV) speech processing in HL populations. Some individuals with HL enhance their perception of spoken language by relying on visual cues (e.g., lip-reading). In a behavioral study, Kaiser et al. (2003) contrasted the ability to identify spoken words in healthy-hearing individuals and in individuals with CIs. They found that all subjects performed better in the AV task, followed by the auditory-only and finally the visual-only task. Also, the healthy-hearing participants outperformed the CI users on the auditory-only condition, while both groups performed at roughly the same level on the AV condition (Kaiser et al., 2003). Furthermore, healthy-hearing musicians and non-musicians seem to process AV speech differently. Recent findings in normal-hearing populations have shown that pitch encoding at the brain stem and auditory cortex is enhanced in musicians compared to non-musicians for AV speech (Musacchia et al., 2007, 2008). FFR responses and wave d (~18 ms) of the ABR were larger for musicians than for non-musicians for speech stimuli presented in an AV context, and FFR enhancement correlated with years of musical training (Musacchia et al., 2007). Also, the cortical P1-N1-P2-N2 complex was larger in musicians than in non-musicians for the consonant-vowel/da/presented in an AV context. Note that enhancements of all the components within the P1-N1-P2-N2 complex have been previously shown to represent manifestations of musical training (Pantev et al., 1998; Schneider et al., 2002; Shahin

et al., 2003; Fujioka et al., 2006). These ABR and cortical enhancements in musicians may reflect superior sensitivity to onsets of speech sounds when modulated by meaningful visual cues, as suggested by Musacchia et al. (2007). Heightened neural sensitivity to sound onsets would improve the identification of word boundaries, a crucial process for word segmentation and thus particularly useful for individuals with HL. Multisensory musical training in a HL population (i.e., watching the instructor's hands during a piano lesson), may not only promote neuroplastic changes in the auditory cortex, but also enhance neurophysiological mechanisms underlying multisensory integration.

Finally, one could ask why musical training? For example, would CI users develop enhanced pitch identification skills following speech discrimination training (Reinke et al., 2003) in the same way following musical training (Chen et al., 2011)? One advantage is that musical training, or even passive listening to music, can provide an emotional experience (Jancke, 2008) unmatched by other acoustical training methods. In a targeted auditory training program, CI users might process sound features more efficiently when enjoying the task, for example, by selecting his/her preferred genre of music (Looi and She, 2010). Targeted auditory training programs have been shown to significantly improve CI users' level of speech recognition (Fu and Galvin, 2008).

GENERAL DISCUSSION

Thus far, I have provided evidence supporting the idea that enhanced basic acoustic (e.g., frequency or rhythm) processing in the brain stem (FFR) and NPAC (P2 AEP) following musical training may be transferable to speech processing. After developing this concept, I then discussed the possibility of applying musical training to individuals with HL to improve their speech perception. These possibilities are exciting, but some limitations of the qualifications of the presented concepts and additional considerations must be discussed here.

The vast majority of studies associating neuroplasticity with musical training have used the musician's brain as a model (Jancke, 2002). A shortcoming of this approach is that distinguished brain attributes seen in musicians may be due to a genetic predisposition, rather than neuroplastic adaptation. Several lines of evidence indicate that this genetic hypothesis is not sufficient to explain these neuroplastic changes observed in musicians. First, enhanced brain responses in musicians are correlated with age of commencement of music lessons, reinforcing the view that musical training is indeed a factor in neuroplastic modifications (Pantev et al., 1998; Schneider et al., 2002; Wong et al., 2007; Lee et al., 2009). Second, short-term acoustical training has been shown to induce neuroplastic effects in the same neural responses that typically distinguish musicians from non-musicians (N1m: Pantev et al., 1998; P2: Shahin et al., 2003). For example, participants showed increased N1m and P2 amplitudes after short-term frequency discrimination training (Menning et al., 2000; Bosnyak et al., 2004). Third, enhanced brain responses in musicians have also been found to be specific for the instrument of practice, implying a neuroplastic effect (Pantev et al., 2001; Neuhaus et al., 2006; Shahin et al., 2008). Finally, several longitudinal studies comparing brain responses in children before and after musical training further underscore the role of musical training on brain plasticity (Fujioka et al., 2006; Shahin

et al., 2008; Hyde et al., 2009; Moreno et al., 2009), with one study providing inconclusive results (Shahin et al., 2004). In Shahin et al. (2004), musically trained children showed larger AEP responses than non-musically trained children before and after commencement of musical training. However, other studies showed enlarged late AEFs/AEPs (N250/N300) in children who took music lessons relative to non-trained children after 1 year (Fujioka et al., 2006), or to art-trained children after 6 months (Moreno et al., 2009) of training. In the study by Moreno et al. (2009), children were randomly assigned to music or art lessons, further strengthening the argument that musical training rather than a genetic predisposition is the primary agent of neuroplastic changes seen in musicians.

Assessing transfer of auditory function following musical training may not be limited to evidence of neuroplastic modifications along the sensory pathway. Processing syntactic violations of musical or speech segments, organizing percepts in working memory, and selecting relevant percepts require the recruitment of higher-level cognitive processes. Syntactic violation processing in music is linked to the early right anterior negativity (ERAN) localized to the left (Broca's region) and right inferior frontal gyri (Maess et al., 2001). Broca's region is also associated with detecting linguistic syntactic violations (Friederici, 2002). The ERAN and ELAN (left-hemisphere speech counterpart) were shown to be larger, and thus more developed, in musically trained children compared to non-trained children following detection of syntactic irregularities in music and speech (Jentschke and Koelsch, 2009). Musically trained children also tended to perform better (with faster reaction time and increased accuracy) than untrained children. Hence, shared mechanisms, as well as a transfer effect, for syntactic violation processing following musical training can be inferred.

It is worth noting that some higher-level auditory functions, such as those associated with the mismatch negativity (MMN), may be specialized for either music or speech processing (Tervaniemi et al., 2000; Tervaniemi and Huutilainen, 2003) and thus may be less susceptible, than FFR and P2 for example, to inter-domain transfer following musical training. The MMN (150–250 ms) is a pre-attentive cognitive process, which has likely sources in A1 and NPAC, as well as inferior frontal cortex (Tervaniemi et al., 2001; Hall, 2006), used to assess the detection of deviancy in an otherwise predictable sequence of events encoded in sensory memory (Naatanen et al., 2007). A larger MMN in musicians than in non-musicians may indicate stronger sensory representations and/or a better ability to predict the next sound in a patterned musical sequence (Koelsch et al., 1999; Fujioka et al., 2004). Currently, the evidence supporting functional overlap for deviant-detection processes associated with the MMN between speech and music domains and the influence of musical training on such overlap are inconclusive (Tervaniemi et al., 1999; Lidji et al., 2009, 2010). A recent study compared the MMN response dynamics for pure tone, music, and speech stimuli in musicians and non-musicians and found that the MMN occurred earlier in musicians than in non-musicians for all types of stimuli (Nikjeh et al., 2009). This suggests that musical training may lead to faster pre-attentive deviant detection, even for speech stimuli. However, there were no group differences in MMN amplitude for any of the stimuli. This is consistent with Tervaniemi et al. (2009) who similarly failed to find amplitude differences in the MMN between musicians and non-musicians for speech stimuli with

deviants in either frequency or duration. However, the MMN was larger in musicians than in non-musicians when participants paid attention to the speech sounds (Tervaniemi et al., 2009). These two studies (Nikjeh et al., 2009; Tervaniemi et al., 2009) suggest that while the MMN mechanisms (at least when considering its amplitude dynamics) may be specialized for music and speech processing, musical training can shape attentional processes (see paragraph below) that could facilitate functional inter-domain transfer for these same neural mechanisms.

One important consideration when assessing transfer effects involves how musical training shapes selective attention mechanisms, a necessary process in noisy or crowded auditory scene situations (Parbery-Clark et al., 2009b; Kerlin et al., 2010). Enhanced selective attention abilities may: (1) impact how relevant and irrelevant signals are organized in working memory (Sreenivasan and Jha, 2007), and (2) act as an auditory gain function in which the relevant acoustical signal intensity is dialed up and the interfering noise/talker is dialed down (Kerlin et al., 2010). The acoustic feature selectively attended to during training shapes cortical map specificity to sound features in A1 (Polley et al., 2006; Engineer et al., 2008). For example, rats trained to attend to frequency cues in sounds showed modifications of the cortical map in A1 corresponding to the trained frequency range only, while rats trained to attend to intensity cues showed modifications of the cortical map to the target intensity range only (Polley et al., 2006). Similarly, during musical training, trainees must focus on or direct attention to slight changes in pitch, intensity, and onsets/offsets of sounds, developing acoustical acuity in the temporal and spectral domains (Schneider et al., 2002; Marie et al., 2010). In turn, enhanced acoustical representations facilitate auditory object formation (acoustical feature binding) and thus allow for better selection and analysis of the acoustic scene (Shinn-Cunningham and Best, 2008). In other words, musicians' improved acuity to sound features, resulting in more veridical auditory object representations, may also shape their selective attention mechanisms, and thus improve their auditory scene analysis skills. Indeed, Strait et al. (2010) revealed that musicians perform better (have faster reaction times) than non-musicians in tasks requiring focused attention in the auditory but not in the visual modality. Additionally, attentive listening to music recruits general cognitive functions related to working memory, semantic processing, and target detection (Janata et al., 2002), which are all processes that influence auditory scene analysis. In sum, enhanced selective attention and other

top-down mechanisms that result from musical training may lead to superior concurrent sound segregation, which is an essential process during perception of speech in noisy environments, especially in individuals with HL.

CONCLUSION

While evidence of neuroplastic adaptations due to musical training is accumulating, the extent to which these neurophysiological changes transfer to speech perception remains inconclusive. Preliminary evidence suggests that musical training may influence a shared hierarchical auditory network underlying music and speech processing and thus can influence speech perception. Further studies are necessary to examine or clarify functional transfer due to musical training along different levels of the auditory pathway (e.g., MLR, MMN) and whether potential transfer effects are solely due to the overlap in acoustical features or also to the overlap in mechanisms that binds them. A possible limitation of this research is that much of the supporting evidence has come from the study of neuroplasticity in musicians, where genetically or ontogenetically determined acoustical abilities could, in principle, have influenced the decision to train musically. However, the dependence of enhanced brain responses in musicians on the duration of training and specific musical experience in addition to evidence provided by longitudinal studies, point to neuroplasticity as the crucial mechanism. Furthermore, future research addressing behavioral and neurophysiological influences of musical training on speech perception in adverse acoustical environments and in individuals with HL is necessary and timely. These topics go hand-in-hand, as speech in background noise creates a debilitating acoustical experience in individuals with HL. Initial evidence shows that musical training enhances pitch perception in individuals with HL. Moreover, musical training is correlated with enhanced behavioral and neurophysiological responses for speech in noise. These exciting preliminary findings provide a stepping stone toward future studies addressing the neurophysiological effects of musical training on speech perception in individuals with HL, which may influence our approach to devising targeted auditory training programs and thus prevention and intervention strategies for HL.

ACKNOWLEDGMENTS

The author thanks Kristina Backer, Dr. Larry Roberts, and Dr. Susan Nittrouer for their helpful comments.

REFERENCES

- Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* 436, 1161–1165.
- Besson, M., and Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278–1296.
- Bever, T. G., and Chiarello, R. J. (2009). Cerebral dominance in musicians and nonmusicians. 1974. *J. Neuropsychiatry Clin. Neurosci.* 21, 94–97.
- Bosnyak, D. J., Eaton, R. A., and Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cereb. Cortex* 14, 1088–1099.
- Chan, A. S., Ho, Y. C., and Cheung, M. C. (1998). Music training improves verbal memory. *Nature* 396, 128.
- Chen, J. K., Chuang, A. Y., McMahon, C., Hsieh, J. C., Tung, T. H., and Li, L. P. (2011). Music training improves pitch perception in prelingually deafened children with cochlear implants. *Pediatrics*, 125: e793–800.
- Engineer, C. T., Perez, C. A., Chen, Y. H., Carraway, R. S., Reed, A. C., Shetake, J. A., Jakkamsetti, V., Chang, K. Q., and Kilgard, M. P. (2008). Cortical activity patterns predict speech discrimination ability. *Nat. Neurosci.* 11, 603–608.
- Freeman, J. A., and Nicholson, C. (1975). Experimental optimization of current source-density technique for anuran cerebellum. *J. Neurophysiol.* 38, 369–382.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci. (Regul. Ed.)* 6, 78–84.
- Fu, Q. J., and Galvin, J. J. III. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hear. Res.* 242, 198–208.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development

- of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608.
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.* 16, 1010–1021.
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J. Cogn. Neurosci.* 17, 1578–1592.
- Gfeller, K., and Lansing, C. R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *J. Speech Hear. Res.* 34, 916–920.
- Hackett, T. A., Preuss, T. M., and Kaas, J. H. (2001). Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J. Comp. Neurol.* 441, 197–222.
- Hall, J. W. (2006). *New Handbook of Auditory Evoked Responses*. Pearson, Boston.
- Hamalainen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Hertrich, I., Mathiak, K., Lutzenberger, W., and Ackermann, H. (2000). Differential impact of periodic and aperiodic speech-like acoustic signals on magnetic M50/M100 fields. *Neuroreport* 11, 4017–4020.
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., and Schlaug, G. (2009). Musical training shapes structural brain development. *J. Neurosci.* 29, 3019–3025.
- Janata, P., Tillmann, B., and Bharucha, J. J. (2002). Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cogn. Affect. Behav. Neurosci.* 2, 121–140.
- Jancke, L. (2002). What is special about the brains of musicians? *Neuroreport* 13, 741–742.
- Jancke, L. (2008). Music, memory and emotion. *J. Biol.* 7, 21.
- Jentschke, S., and Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage* 47, 735–744.
- Kaiser, A. R., Kirk, K. I., Lachs, L., and Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J. Speech Lang Hear Res.* 46, 390–404.
- Kerlin, J. R., Shahin, A. J., and Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *J. Neurosci.* 30, 620–628.
- Koelsch, S., Schroger, E., and Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport* 10, 1309–1313.
- Krishnan, A., Swaminathan, J., and Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J. Cogn. Neurosci.* 21, 1092–1105.
- Lee, K. M., Skoe, E., Kraus, N., and Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.* 29, 5832–5840.
- Lenz, D., Jeschke, M., Schadow, J., Naue, N., Ohl, F. W., and Herrmann, C. S. (2008). Human EEG very high frequency oscillations reflect the number of matches with a template in auditory short-term memory. *Brain Res.* 1220, 81–92.
- Lidji, P., Jolicoeur, P., Kolinsky, R., Moreau, P., Connolly, J. F., and Peretz, I. (2010). Early integration of vowel and pitch processing: a mismatch negativity study. *Clin. Neurophysiol.* 121, 533–541.
- Lidji, P., Jolicoeur, P., Moreau, P., Kolinsky, R., and Peretz, I. (2009). Integrated preattentive processing of vowel and pitch: a mismatch negativity study. *Ann. N. Y. Acad. Sci.* 1169, 481–484.
- Looi, V., and She, J. (2010). Music perception of cochlear implant users: a questionnaire, and its implications for a music training program. *Int. J. Audiol.* 49, 116–128.
- Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: a MEG study. *Nat. Neurosci.* 4, 540–545.
- Marie, C., Magne, C., and Besson, M. (2010). Musicians and the metric structure of words. *J. Cogn. Neurosci.* 23, 294–305.
- Menning, H., Roberts, L. E., and Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport* 11, 817–822.
- Meyer, M., Baumann, S., and Jancke, L. (2006). Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage* 32, 1510–1523.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., and Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* 19, 712–723.
- Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894–15898.
- Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.* 241, 34–42.
- Naatanen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590.
- Neuhaus, C., and Knösche, T. R. (2008). Processing of pitch and time sequences in music. *Neurosci. Lett.* 441, 11–15.
- Neuhaus, C., Knösche, T. R., and Friederici, A. D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *J. Cogn. Neurosci.* 18, 472–493.
- Nikjeh, D. A., Lister, J. J., and Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training. *Ear Hear.* 30, 432–446.
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., and Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature* 392, 811–814.
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., and Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12, 169–174.
- Parbery-Clark, A., Skoe, E., and Kraus, N. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29, 14100–14107.
- Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear Hear.* 30, 653–661.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–776.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., and Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117(Pt 6), 1283–1301.
- Polley, D. B., Steinberg, E. E., and Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* 26, 4970–4982.
- Preisler, A. (1993). The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch. *Percept. Psychophys.* 54, 589–603.
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114.
- Reinke, K. S., He, Y., Wang, C., and Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Brain Res. Cogn. Brain Res.* 17, 781–791.
- Riecke, L., Esposito, F., Bonte, M., and Formisano, E. (2009). Hearing illusory sounds in noise: the timing of sensory-perceptual transformations in auditory cortex. *Neuron* 64, 550–561.
- Rogalsky, C., Rong, F., Saberi, K., and Hickok, G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J. Neurosci.* 31, 3843–3852.
- Sammler, D., Koelsch, S., Ball, T., Brandt, A., Elger, C. E., Friederici, A. D., Grigutsch, M., Huppertz, H. J., Knosche, T. R., Wellmer, J., Widman, G., and Schulze-Bonhage, A. (2009). Overlap of musical and linguistic syntax processing: intracranial ERP evidence. *Ann. N. Y. Acad. Sci.* 1169, 494–498.
- Scherg, M. (1990). “Fundamentals of dipole source potential analysis,” in *Auditory Evoked Magnetic Fields and Electric Potentials*, Vol. 6, *Advances in Audiology*, eds F. Grandori, M. Hoke, and G. L. Romani (Basel: Karger), 40–69.
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., and Rupp, A. (2002). Morphology of Heschl’s gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694.
- Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., and Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in non-musicians and musicians. *Hum. Brain Mapp.* 32, 771–783.
- Seither-Preisler, A., Johnson, L., Krumbholz, K., Nobbe, A., Patterson, R., Seither, S., and Lutkenhoner, B. (2007). Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 743–751.
- Seither-Preisler, A., Krumbholz, K., and Lutkenhoner, B. (2003). Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? *Audiol. Neurootol.* 8, 322–337.
- Shahin, A., Bosnyak, D. J., Trainor, L. J., and Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory

- evoked potentials in musicians. *J. Neurosci.* 23, 5545–5552.
- Shahin, A., Roberts, L. E., and Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport* 15, 1917–1921.
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., and Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport* 16, 1781–1785.
- Shahin, A. J., Bishop, C. W., and Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44, 1133–1143.
- Shahin, A. J., Roberts, L. E., Chau, W., Trainor, L. J., and Miller, L. M. (2008). Music training leads to the development of timbre-specific gamma band activity. *Neuroimage* 41, 113–122.
- Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299.
- Snyder, J. S., and Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Res. Cogn. Brain Res.* 24, 117–126.
- Song, J. H., Skoe, E., Wong, P. C., and Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cogn. Neurosci.* 20, 1892–1902.
- Sreenivasan, K. K., and Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *J. Cogn. Neurosci.* 19, 32–41.
- Strait, D. L., Kraus, N., Parbery-Clark, A., and Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear. Res.* 261, 22–29.
- Tervaniemi, M., and Huottilainen, M. (2003). The promises of change-related brain potentials in cognitive neuroscience of music. *Ann. N. Y. Acad. Sci.* 999, 29–39.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., and Schroger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp. Brain Res.* 161, 1–10.
- Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., and Friederici, A. D. (2009). Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *Eur. J. Neurosci.* 30, 1636–1642.
- Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., and Naatanen, R. (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *Neuroimage* 9, 330–336.
- Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., and Naatanen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10, 74–79.
- Tervaniemi, M., Rytönen, M., Schroger, E., Ilmoniemi, R. J., and Naatanen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learn. Mem.* 8, 295–300.
- Tervaniemi, M., Szameitat, A. J., Kruck, S., Schroger, E., Alter, K., De Baene, W., and Friederici, A. D. (2006). From air oscillations to music and speech: functional magnetic resonance imaging evidence for fine-tuned neural networks in audition. *J. Neurosci.* 26, 8647–8652.
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., and Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear.* 22, 79–90.
- Tremblay, K. L., and Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *J. Speech Lang. Hear. Res.* 45, 564–572.
- Tremblay, K. L., Shahin, A. J., Picton, T., and Ross, B. (2009). Auditory training alters the physiological detection of stimulus-specific cues in humans. *Clin. Neurophysiol.* 120, 128–135.
- Wang, W., Zhou, N., and Xu, L. (2011). Musical pitch and lexical tone perception with cochlear implants. *Int. J. Audiol.* 50, 270–278.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392–393.
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422.
- Yucel, E., Sennaroglu, G., and Belgin, E. (2009). The family oriented musical training for children with cochlear implants: speech and musical perception results of two year follow-up. *Int. J. Pediatr. Otorhinolaryngol.* 73, 1043–1052.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci. (Regul. Ed.)* 6, 37–46.
- Zendel, B. R., and Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *J. Cogn. Neurosci.* 21, 1488–1498.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 February 2011; accepted: 30 May 2011; published online: 13 June 2011.

Citation: Shahin AJ (2011) Neurophysiological influence of musical training on speech perception. *Front. Psychology* 2:126. doi: 10.3389/fpsyg.2011.00126

This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*. Copyright © 2011 Shahin. This is an open-access article subject to a non-exclusive license between the authors and *Frontiers Media SA*, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other *Frontiers* conditions are complied with.