Critical Review: Do individuals with musical training have enhanced brainstem encoding of linguistic pitch compared to those not musically trained?

Kryszak, J.K.
M.Cl.Sc. (Aud) Candidate
School of Communication Sciences and Disorders, U.W.O

This critical review examines the effects of musical training on brainstem encoding of linguistic pitch. Furthermore, it investigates whether individuals with musical training have enhanced brainstem encoding of linguistic pitch when compared to those not musically trained. Overall, research suggests more robust and faithful encoding of linguistic pitch information by musicians. Musicians also display an enhanced representation of the fundamental frequency, which has been extensively understood to underlie pitch perception. Thus, the two studies reviewed indicate that musicians have better brainstem encoding of linguistic pitch, which suggests a context-general corticofugal tuning of the afferent system. These findings have implications for general social and educational policies with regards to the value of musical training in schools and auditory training strategies for individuals with speech-encoding deficits. However, due to the recent development of this research more thorough and systematic investigations of musician and non-musicians' responses to different simple and complex sounds are necessary.

Introduction

Music and speech are two cognitively demanding auditory stimuli that are often attributed to cortical rather than subcortical circuitry (Wong et al., 2007). Previous research confirms that musical training modifies cortical organization (Musacchia et al., 2007). However, less is known about how long-term complex sound experiences such as music, shape subcortical circuitry (Wong et al., 2007).

It has long been established that recording brainstem responses to sound is a valid and reliable means to assess the integrity of the neural transmission of acoustic stimuli (Johnson et al., 2005). The human auditory brainstem response accurately reflects the frequency and time-varying characteristics of sound and has been studied extensively using click, tonal and speech stimuli (Musacchia et al., 2007).

In the case where the brainstem response is elicited by a speech syllable, the response itself can be divided into transient and sustained portions. The transient response to speech onset is comparable to that of a click-evoked response (Musacchia et al., 2007). The sustained portion is called the frequency-following response (FFR) and arises from the harmonic portion of the stimulus. It is characterized as a series of transient neural events phase-locked to periodic information within the stimulus (Johnson et al., 2005).

Since the FFR is presumed to originate from the auditory brainstem (inferior colliculus) and is able to encode the energy of the stimulus fundamental frequency (f0) with high fidelity, researchers are able to use it as a means of observing linguistic pitch pattern at the level of the brainstem (Wong et al., 2007).

Previous work completed by Krishnan et al. (2005) found increased linguistic pitch pattern encoding in Mandarin-speaking subjects relative to English speaking subjects. The results obtained reflect the Mandarin-speaking subjects' long-term familiarity with linguistic pitch patterns as Mandarin Chinese is a tonal language that uses pitch to signal word meaning. Krishnan et al. (2005) suggest that these results support the likelihood of neural plasticity at the level of the brainstem, which is stimulated by language experience and therefore may be responsible for enhancing or priming linguistically relevant features of the speech input.

The following review will examine the effect of music-related experience and whether musical training is able to enhance brainstem encoding of linguistic pitch.

Objectives

The primary objective of this paper is to critically evaluate existing literature regarding the impact of musical training on brainstem encoding of linguistic pitch. The secondary objective is to generate implications of the findings as well as provide future directions for research.

Methods

Search Strategy

Computerized databases, including MEDLINE-OVID and PubMed were searched using the following search strategy:

((Musicians) OR (musical training)) AND (brainstem encoding)

No limitations were applied to these searches.

Selection Criteria

Studies selected for inclusion in this critical review paper were required to investigate the impact of musical training on brainstem encoding of linguistic pitch. No limits were set on the demographics of research participants or outcome measures.

Data Collection

Results of the literature search yielded two articles congruent with the aforementioned selection criteria: case control (2). The first study did not discuss participant selection, experimenter or performance bias. However, the two subject groups were well matched and the study did account for measurement artifact, order effect and environmental contexts. The second study did not discuss participant selection or performance bias. However, the study did account for experimenter bias, measurement artifact and order effect.

Results

The first study by Wong et al. (2007) measured the frequency following response (FFR) to linguistic pitch patterns at the rostral brainstem in both amateur musicians and non-musicians. The second study by Musacchia et al. (2007) used temporal and spectral resolution of the auditory brainstem response to examine whether, and to what extent, subcortical processing could be shaped by musical experience.

Wong et al. (2007)

This case study involved 20 subjects that were equally divided into two groups based on musical training. Amateur musicians had at least 6 years of continuous musical training before the age of 12 and were currently playing at the time of the study. The non-musicians had no more than 3 years of musical training. All subjects were right handed, had normal thresholds at 20dB or below and normal click evoked auditory brainstem responses (ABR). The two subject groups did not differ in age.

Subjects listened to a video sound track set at <40dB SPL in a free field. The left ear was unoccluded while three Mandarin stimuli were randomly presented to the right ear at 70 dB SPL. The three Mandarin stimuli which were resynthesized to differ only in f0: m1 “squint”, m2 “bewilder”, m3 “rice”. The number given to each stimulus indicated tone or lexically meaningful pitch contour: Tone 1 – level tone, Tone 2 – rising tone and Tone 3- dipping tone.

Brainstem responses were collected using Scan 4.3 with Ag-AgCl scalp electrodes.

Researchers derived to two primary pitch-tracking measures. The first was a stimulus –to-response correlation used to determine the faithfulness of pitch tracking. The second was a peak auto-correlation used to indicate the robustness of neural phase-locking.

The primary measures were entered into a 3 (tone) x 2 (group) repeated measures ANOVA. The stimulus-to-response correlation revealed a main effect of group, $P < 0.015$, and tone, $P <0.001$, but no significant interaction. It was concluded that musicians showed more precise pitch-tracking compared to non-musicians.

For peak auto-correlation there was a significant effect of tone, $P<0.001$, but not of group and a marginally significant interaction, $P <0.08$. Visual inspection of the autocorrelation plots indicated musicians to have more robust responses for Tone 3 but not for the others. A one independent samples t-test was conducted on Tone 3, which confirmed musicians’ higher autocorrelation values.

Overall, musicians displayed a more faithful representation of the stimulus f0 contours and more robust neural phase-locking, this was especially true for Tone 3, the most complex contour.

Musacchia et al. (2007)

This case study involved 29 subjects that were divided into two groups based on musical experience. The musician group began playing an instrument before the age of 5 and had 10 years or more of experience and practiced 3 times a week for 4 hours or more over the last 10 years. The control group was made up of those individuals who failed to meet the musician’s criteria. All subjects had normal hearing, normal or corrected to normal vision, and no history of neurological disorders.

Six types of stimuli were presented: Unimodal acoustic - UA (speech –syllable ‘da’/musical stimulus –cello being bowed note G2), Unimodal visual -UV (speech-video of a male speaker articulating ‘da’ / musical –video of a musician bowing a cello) and congruent pairing of UA and UV to create an audiovisual (AV) stimulus.

Experimenters controlled for attention by asking subjects to silently count the number of target stimuli they saw or heard and then report that number at the end of each block.

Continuous electroencephalographic data were recorded from Cz (10-20 international system). Brainstem onset response peaks (wave V, A, δ, γ) were chosen from each individual's responses providing latency and amplitude information. Peak latencies were calculated by subtracting the latency of sound onset (time 0) from the latency of the peak voltage fluctuation for each wave. The strength of pitch encoding was measured by peak amplitudes at F0 (100Hz) and from H2 (200Hz) to H5 (500Hz) of the fast Fourier transforms over the FFR period.

Results indicated that musicians had earlier brainstem responses than non-musicians to speech onset in both the UA and AV modalities. This was evident as early as 10ms after acoustic onset.

Group differences were observed in the frequency-following portion of the response. The grand average fast Fourier transform of responses over time for speech indicated that musicians have enhanced periodicity encoding (phase-locking), particularly linked to the f0 and throughout the FFR period. Amplitudes were larger in musicians than controls for both the UA ($t = 2.81$, $P<0.0125$) and AV ($t = 2.72$, $P <0.0125$) conditions.

Speech evoked f0 amplitudes correlated positively with the number of years musicians had been faithfully playing music within the past 10 years. The effect was observed in both the UA ($r=0.731$, $P = 0.001$) and AV ($r = 0.68$, $P<0.01$) conditions.

Overall, musicians had earlier and larger brainstem responses compared to non-musicians when both speech and music stimuli were presented in the auditory and audiovisual conditions. Phase-locking to stimulus periodicity was found to be enhanced in musicians and was strongly correlated with the length of musical training.

Discussion

Musicians have extensive experience using pitch information in the context of music, which involves both high cognitive demands and precise auditory perception (Wong et al., 2007). Previous research has shown that musical talent predicts the ability to produce and perceive the sound structures, but not grammatical or semantic structures. Therefore, musicians seem to have an enhanced ability to learn lexical tones (Slevc & Miyake, 2006).

In the studies discussed above, researchers have found a plausible subcortical correlate of the effect of long-term musical training on prosodic encoding of speech (Wong et al., 2007). The present literature suggests that musical training does enhance the ability of the brainstem to encode linguistic pitch. The two studies reviewed here have shown that musicians have a more faithful representation of linguistic pitch as well as more robust neural phase-locking when elicited by a non musical stimulus. These findings suggest that corticofugal modulation is not entirely context specific (Wong et al., 2007). The earlier latencies and larger onset responses displayed by musicians imply that these individuals have a more synchronous neural response to the onset of sound, which is characteristic of a high functioning peripheral auditory system (Musacchia et al., 2007). Musicians also show a superior representation of the f0, which is extensively understood to underlie pitch perception. Data also indicates that intensive musical practice and exposure relate to the strength of pitch encoding (Musacchia et al., 2007).

The aforementioned studies provide evidence for the positive effect of long-term music exposure on linguistic pitch encoding at the level of brainstem but the data obtained from these studies cannot accurately depict which aspect of musical training is responsible for the enhanced subcortical encoding. Musical training involves discrimination of pitch intonation, onset, offset, and duration aspects of sound timing as well as integration of multisensory cues (Musacchia et al., 2007). Therefore, a variety of elements may be responsible for the enhancement of brainstem encoding seen in those individuals with musical training.

Alternatively, the difference between these two groups may be due to the fact that musicians have learned to pay more attention to the details of acoustic stimuli (Musacchia et al., 2007). Also genetic differences between musicians and non-musicians may potentially account for the difference in results (Wong et al., 2007).

Conclusion

Outcomes from the studies indicate more robust and faithful encoding of linguistic pitch by musicians when compared to those not musically trained. The enhancement in linguistic encoding seen by musicians is arguably related to the increase in musical pitch usage and may reflect a positive side effect of context-general corticofugal tuning of the afferent system. This would imply that long-term musical training might have a role in shaping basic sensory circuitry (Wong et al., 2007). These findings are in agreement with existing knowledge of the brainstem's role in encoding speech (Johnson et al., 2005) by illustrating the interaction between music and speech, subcortical and cortical structures, and the effect of long-term auditory experiences (Wong et al., 2007).

Possible Implications and Future Directions For Research

The findings from the reviewed articles have practical implications for general social and educational policies (Wong et al., 2007). These implications are especially important when considering the value of musical training in schools and the investigation of auditory training strategies for individuals with speech-encoding deficits (Musacchia et al., 2007).

The aforementioned studies suggest that high-level, complex training, such as learning to play a musical instrument, effects encoding mechanisms in peripheral sensory structures (Musacchia et al., 2007). However, it is difficult to accurately depict which aspect of musical training is responsible for the enhanced subcortical encoding. (Wong et al., 2007).

Future research should take a thorough and systematic approach to investigate musician and non-musicians' responses to various simple and complex sounds (Wong et al., 2007). Also research focusing on the minimum amount of musical training needed to observe the positive effects of enhanced linguistic encoding would be valuable. Finally, studies concerning the mechanisms attributed to the enhanced linguistic encoding in musicians should be continued with the specific goal of determining how and where these changes are taking place.

References

Johnson, K.L., Nicol, T.G., & Kraus, N. (2005). Brain stem response to speech: a biological marker of auditory processing. *Ear and Hearing, 5*, 424-434.

Krishnan A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25,* 161-168.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 15894-15898.

Slevc., D.L., & Miyake, A. (2006). Individual differences in second-language proficiency: does music ability matter? *Psychological Science, 17,* 675-681.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, N., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10,* 420-422.