Musical Experience Promotes Subcortical Efficiency in Processing Emotional Vocal Sounds

Dana L. Strait,[a,b] Nina Kraus,[b,c,d,e] Erika Skoe,[b,c] and Richard Ashley[a,f]

[a]Bienen School of Music, [b]Auditory Neuroscience Laboratory, Departments of [c]Communication Sciences, [d]Neurobiology and Physiology, [e]Otolaryngology, and [f]Cognitive Science at Northwestern University, Evanston, Illinois, USA

To understand how musical experience influences subcortical processing of emotionally salient sounds, we recorded brain stem potentials to affective vocal sounds. Our results suggest that auditory expertise engenders subcortical auditory processing efficiency that is intricately connected with acoustic features important for the communication of emotion. This establishes a subcortical role in the auditory processing of emotional cues, providing the first biological evidence for musicians' enhanced perception of vocally expressed emotion.

Key words: **auditory brain stem response; brain; emotion; musicians; plasticity**

Perception of emotion in speech and music relies on shared acoustic and neural mechanisms,[1] suggesting that extensive experience in one domain may lend perceptual benefits to the other. Accordingly, musical experience enhances perceptual sensitivity to emotion in speech.[2,3]

Musical experience also shapes subcortical sound transcription (see the paper by Kraus *et al.* in this volume[4]). Because of its fidelity in representing spectral and temporal acoustic features, the auditory brainstem response (ABR) provides a mechanism for exploring musicians' subcortical sensitivity to acoustic features contributing to the perception of emotion.[5] ABRs may also advance our understanding of subcortical function in the processing of emotionally charged auditory cues.

To better understand musical experience's influence on neural processing of affective speech-related events, we recorded ABRs to an emotionally charged vocal sound—an infant's unhappy cry. We aimed to provide a biological basis for musicians' enhanced perception of emotion in speech by investigating the contribution of subcortical mechanisms to processing vocally communicated emotion.

Methods

Subjects were 30 normal-hearing adults aged 19–35 years, with musicians grouped according to two criteria: "Musicians by Onset Age" (MusAge, $n = 11$, began musical training ≤ 7 yr) and "Musicians by Years" (MusYrs, $n = 15$, ≥ 10 years of consistent musical training). Nonmusicians (NonMus) were categorized by failure to meet those criteria.

ABRs were elicited by a complex vocal sound derived from an emotional auditory scene from the Center for the Study of Emotions and Attention (University of Florida, Gainesville; file 278).[6] For further recording and data processing parameters see Strait *et al.*[7]

Address for correspondence: Dana L. Strait, Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, IL 60208. Voice: 847-491-2465. Dana.Strait@u.northwestern.edu

[b]For more information about the Auditory Neuroscience Laboratory and the work presented herein, please visit http://www.brainvolts.northwestern.edu.
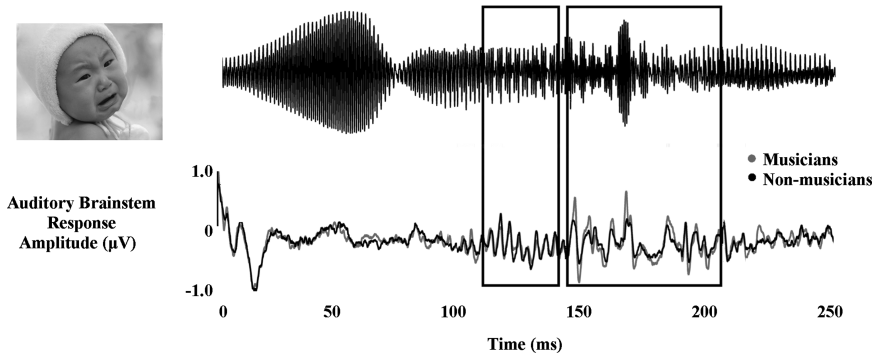
Figure 1. Stimulus and grand average response waveforms. The boxes correspond to the periodic and complex portions, respectively. Figures adapted from Strait *et al.*[7] with permission. (In color in *Annals* online.)

We divided the stimulus into two segments ("periodic" and "complex") that were acoustically contrastive and internally consistent. Neural responses to the complex region resulted in a series of peaks that aligned with amplitude bursts in the stimulus (Fig. 1). Peak latencies and amplitudes were recorded for the largest and most replicable peaks, whereas rectified mean amplitudes (RMAs) provided a gross measure of response amplitude.

We extracted spectral components of the neural responses using the fast Fourier transform. Amplitudes were recorded for spectral peaks corresponding to the stimulus's fundamental frequency and spectral components representing the maxima within given frequency ranges (F_0: 280–305; H_2: 470–485; H_3: 570–585 Hz), interpreted to correspond to representations of pitch (F_0) and timbre (H_2, H_3).

Results

Regression analyses supported the grouping of musicians into two subgroups: whereas the MusYrs grouping was predicted best by amplitude and latency measures ($P < 0.001$; MusAge $P < 0.06$), MusAge was predicted best by frequency encoding measures ($P < 0.05$; MusYrs $P < 0.40$).

Both MusYrs and MusAge musicians' responses exhibited "enhancements" and "econ-omy" connected with time-varying acoustic features of the stimulus. Enhancements (larger time- and frequency-domain response magnitudes) were apparent in musicians' responses to the complex portion of the sound, with economy (smaller amplitudes) seen in their responses to the periodic portion.

Figure 1 shows both the stimulus and average responses for MusYrs and NonMus, with boxes defining acoustically distinct sections. The first portion of the stimulus is characterized by greater periodicity, whereas the second is characterized by greater complexity (deviation in pitch from the F_0: 0.5% in the periodic portion and 2.7% in the complex, harmonic jitter: 1.27% and 2.03%, and signal-to-noise ratio: 13.46 and 6.82 dB).

The response RMAs are plotted in Figure 2A, for which an ANOVA revealed an interaction between group and response portion ($F = 6.04$, $P < 0.02$). This indicates that MusYrs and NonMus have differentiated responses to the two sections. Whereas the MusYrs within-group RMAs differ between the periodic and complex portions ($t = 4.70$, $P < 0.0001$), NonMus do not ($t = 0.025$, $P < 0.99$). Peak amplitudes confirm MusYrs to have larger responses to the complex portion than NonMus (peak 1: $F = 10.25$, $P < 0.003$; peak 2: $F = 4.88$, $P < 0.03$).

Peak amplitudes within the complex portion correlated with years of musical practice
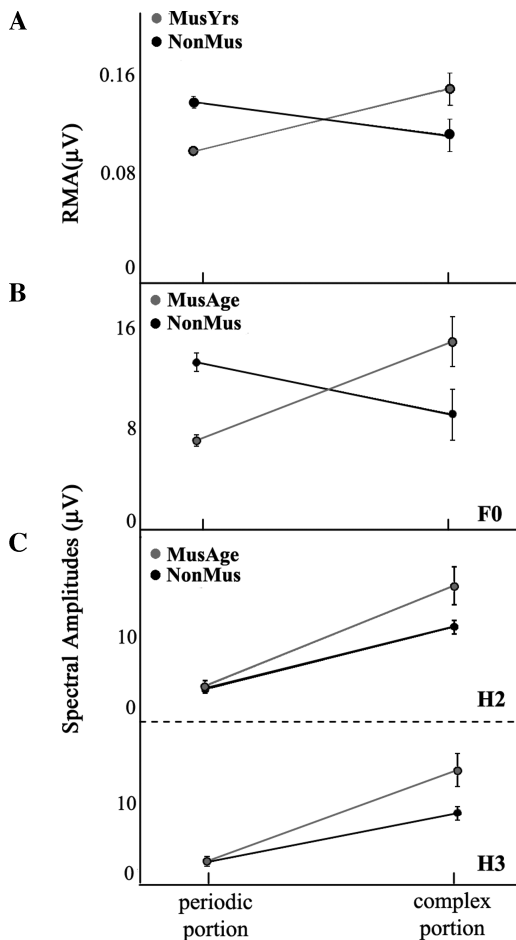
Figure 2. Interactions between responses to the periodic and complex stimulus portions. (**A**) Group × portion interaction between MusYrs and NonMus response RMAs and (**B**) MusAge and NonMus F_0. (**C**) MusAge also show enhanced encoding of frequencies above the F_0 in responses to the complex portion. (In color in *Annals* online.)

across all individuals with musical experience ($n = 20$; Fig. 3; peak 1: $r = 0.454$, $P < 0.04$). Timing-related enhancements were specifically observed in MusYrs responses, even as early as the onset (Fig. 4; onset peak: $F = 4.82$, $P < 0.04$; peak 1: $F = 8.72$, $P < 0.006$). Earlier latencies in musicians reflect faster synchronous neural responses to the timing characteristics of the stimulus.

Compared to NonMus, MusAge showed enhanced representations of frequencies important for the perception of pitch and timbre in responses to the complex stimulus portion. These differences were connected with acoustic characteristics of the auditory input. An ANOVA revealed an interaction between group and F_0 encoding for the two response portions ($F = 7.04$, $P < 0.01$), indicating that MusAge and NonMus have differentiated F_0 responses to the two sections. MusAge showed smaller representations of the F_0 in responses to the periodic portion of the stimulus than NonMus, but larger F_0 amplitudes to the complex portion (Fig. 2B; periodic: $F = 7.04$, $P < 0.01$; complex: $F = 5.04$, $P < 0.03$). F_0 amplitudes in responses to the complex portion correlated with the age that musical training began, with individuals who began at an earlier age showing larger F_0 representations (Fig. 3: $r = -0.500$, $P < 0.03$). MusAge also demonstrated enhanced representations of spectral peaks H_2 and H_3 in responses to the complex portion of the stimulus compared with NonMus (Fig. 2C; H_2: $F = 7.95$, $P < 0.01$; H_3: $F = 6.16$, $P < 0.02$).

Conclusions

We suggest that musical experience has pervasive effects on the auditory system, resulting in fine neural tuning to acoustic features important for vocal communication. Musical experience sharpens subcortical auditory processing, with the behavioral relevance and relative complexity of the stimulus playing a prominent role in subcortical malleability. This sharpening is likely mediated by the corticofugal system known to shape receptive field properties in the primary auditory cortex, thalamus, and auditory brain stem.[8]

The interplay between response enhancement and economy may engender musicians' enhanced perceptual capabilities of emotional cues in speech.[2,3] Musicians' responses to the periodic stimulus section demonstrate smaller magnitudes, reflecting neural efficiency (recruitment of fewer resources) in processing simpler acoustic features. Such observations have been interpreted to reflect domain-general expertise.[9] MusYrs' faster responses indicate that
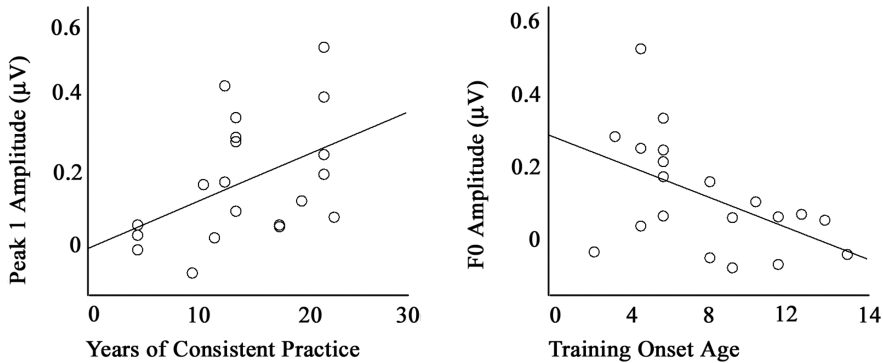
Figure 3. Correlations between musical experience and subcortical response characteristics.
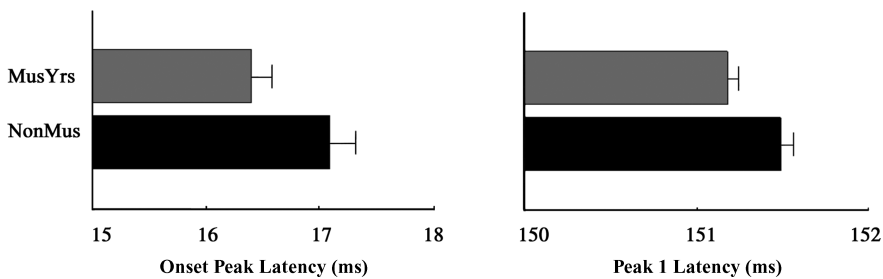

Figure 4. Peak latencies for MusYrs and NonMus. (In color in *Annals* online.)

long-term musical experience contributes to enhanced subcortical timing.

Music's spectral and temporal complexity makes it a powerful tool for engendering neural plasticity during optimal periods of auditory development. Our data suggest that musical training prior to the age of seven has an impact on subcortical frequency representation, whereas timing-related enhancements are affected by duration of practice. This indicates an optimal period for the development of pitch and timbre encoding strategies. Deprivation studies provide evidence for optimal periods in the acquisition of tonotopic maps in the primary auditory cortex, with exposure to spectrally and temporally complex auditory input necessary for auditory evoked-response development.[10,11]

There does not seem to be a similar optimal period for the development of neural representations of timing. The discrepancy between timing- and frequency-related effects aligns with evidence of distinct subcortical encoding mechanisms for different features of acoustic stimuli.[9,12]

Fast subcortical-limbic pathways for emotional processing are established in the visual system,[13] with analogous pathways in the auditory system indicated more recently.[14–16] By showing subcortical involvement in the encoding of acoustic features foundational to the vocal communication of emotion, our results provide an advance in the study of human perception of biological states.

In conclusion, we found that musical training engenders subcortical efficiency that is connected with acoustic features integral to the communication of emotion. Thus we provide biological evidence for musicians' perceptual enhancements in detecting vocally expressed emotion and reveal profound interaction between cognitive and sensory processes.[4,9,16–18]

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Nair, D. *et al*. 2002. Expressive timing and perception of emotion in music: an fMRI study. In *Proceedings of the 7th International Conference on Music Perception and Cognition*. C. Stevens et al., Eds.: 627–630. Causal Productions. Adelaide, Australia.

2. Dmitrieva, E.S. *et al*. 2006. Ontogenetic features of the psychophysiological mechanisms of perception of the emotional component of speech in musically gifted children. *Neurosci. Behav. Physiol.* **36:** 53–62.

3. Thompson, W.F., E.G. Schellenberg & G. Husain. 2004. Decoding speech prosody: do music lessons help? *Emotion* **4:** 46–64.

4. Kraus, N. *et al*. 2009. Experience-induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. *Ann. N. Y. Acad. Sci.* The Neurosciences and Music III: Disorders and Plasticity. **1169:** 543–557.

5. Juslin, P.N. & P. Laukka. 2003. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* **129:** 770–814.

6. Bradley, M.M. & P.J. Lang. 2007. *The International Affective Digitized Sounds (2nd ed.; IADS-2): Affective Ratings of Sounds and Instruction Manual*. NIMH Center for the Study of Emotion and Attention. University of Florida. Gainesville, FL.

7. Strait, D.L. *et al*. 2009. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *Eur. J. Neurosci.* **29:** 661–668.

8. Luo, F. *et al*. 2008. Corticofugal modulation of initial sound processing in the brain. *J. Neurosci.* **28:** 11615–11621.

9. Grabner, R.H., A.C. Neubauer & E. Stern. 2006. Superior performance and neural efficiency: the impact of intelligence and expertise. *Brain Res. Bull.* **69:** 422–439.

10. Trainor, L.J. 2005. Are there critical periods for musical development? *Dev. Psychobiol.* **46:** 262–278.

11. Fallon, J.B., D.R.F. Irvine & R.K. Shepherd. 2002. Cochlear implants and brain plasticity. *Br. Med. Bull.* **66:** 321–330.

12. Kraus, N. & T. Nicol. 2005. Brainstem origins for cortical "what" and "where" pathways in the auditory system. *Trends Neurosci.* **28:** 176–181.

13. Johnson, M.H. 2005. Subcortical face processing. *Nat. Rev. Neurosci.* **6:** 766–774.

14. Bigand, E., S. Filipic & P. Lalitte. 2005. The time course of emotional responses to music. *Ann. N. Y. Acad. Sci.* **1060:** 429–437.

15. Brandao, M.L. *et al*. 2005. Gabaergic regulation of the neural organization of fear in the midbrain tectum. *Neurosci. Biobehav. Rev.* **29:** 1299–311.

16. Marsh, R.A. *et al*. 2002. Projection to the inferior colliculus from the basal nucleus of the amygdala. *J. Neurosci.* **22:** 10449–10460.

17. Banai, K. *et al*. 2009. Reading and subcortical auditory function. *Cereb. Cortex*. In press. doi: 10.1093/cercor/bhp024.

18. Tzounopoulos, T. & N. Kraus. 2009. Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* **62:** 463–469.