

Musical Training Enhances Neural Processing of Binaural Sounds

Alexandra Parbery-Clark,^{1,2} Dana L. Strait,^{1,3} Emily Hittner,^{1,4} and Nina Kraus^{1,2,3,5,6}

¹Auditory Neuroscience Laboratory, ²Institute for Neuroscience, ³Department of Communication Sciences, ⁴Weinberg College of Arts and Sciences, ⁵Department of Neurobiology and Physiology, and ⁶Department of Otolaryngology, Northwestern University, Evanston, Illinois 60208

While hearing in noise is a complex task, even in high levels of noise humans demonstrate remarkable hearing ability. Binaural hearing, which involves the integration and analysis of incoming sounds from both ears, is an important mechanism that promotes hearing in complex listening environments. Analyzing inter-ear differences helps differentiate between sound sources—a key mechanism that facilitates hearing in noise. Even when both ears receive the same input, known as diotic hearing, speech intelligibility in noise is improved. Although musicians have better speech-in-noise perception compared with non-musicians, we do not know to what extent binaural processing contributes to this advantage. Musicians often demonstrate enhanced neural responses to sound, however, which may undergird their speech-in-noise perceptual enhancements. Here, we recorded auditory brainstem responses in young adult musicians and non-musicians to a speech stimulus for which there was no musician advantage when presented monaurally. When presented diotically, musicians demonstrated faster neural timing and greater intertrial response consistency relative to non-musicians. Furthermore, musicians' enhancements to the diotically presented stimulus correlated with speech-in-noise perception. These data provide evidence for musical training's impact on biological processes and suggest binaural processing as a possible contributor to more proficient hearing in noise.

Introduction

The mammalian auditory system adaptably encodes the sounds around us. One way it achieves this is through binaural processing: the ability to simultaneously integrate and analyze input from both ears. The auditory system is tuned to detect minute differences between the ears on the order of tens to hundreds of microseconds (Hudspeth, 1997; Grothe, 2003). Such precision is necessary for hearing in noise, which relies on timing mechanisms to segregate concurrent acoustic streams according to slight deviations in their locations, pitches, or sound qualities. Even when both ears receive identical auditory input (i.e., diotic stimuli that do not contain different level, timing, localization, or pitch cues), hearing thresholds in noise and intelligibility are improved (Plomp and Mimpen, 1979; Kaplan and Pickett, 1981; Davis and Haggard, 1982; Davis et al., 1990). Knowing how auditory experience refines diotic processing may help us better understand this facet of binaural hearing, thus guiding the development of habilitation and remediation approaches for communication abilities for which binaural hearing is a primary concern, including speech perception in noise.

While acoustic input is first relayed ipsilaterally, contralateral projections facilitate the integration of sensory input to both ears early in the central processing stream: binaural interactions are initiated within the superior olivary complex, the nuclei of the lateral lemniscus, and the inferior colliculus of the auditory brainstem (for review, see Moore, 1991). Binaural hearing proficiency does not rely on these subcortical mechanisms alone: during hearing, left and right ear inputs compete within both right and left auditory cortices (Fujiki et al., 2002). Degraded auditory experiences, such as those associated with prior reduced hearing ability (Hogan and Moore, 2003), and cognitive functions, such as attention (Shinn-Cunningham et al., 2005), influence how the inputs from both ears interact. Over the course of their training, musicians develop the ability to make sense of complex auditory environments as well as demonstrating enhanced perceptual learning abilities (Kühnis et al., 2013; Shook et al., 2013). Growing evidence indicates that musical experience is associated with strengthened perception and neural encoding of speech in the presence of noise (Zendel and Alain, 2009; Parbery-Clark et al., 2009a, 2011; Strait et al., 2012, 2013b; Strait and Kraus, 2013), in addition to cognitive abilities that modulate both perception and neural response properties (Parbery-Clark et al., 2009b, 2011; Kraus and Chandrasekaran, 2010; Strait et al., 2010, 2013a; Hanna-Pladdy and MacKay, 2011). Musicians' cognitive and speech-in-noise perceptual benefits have only been measured using binaural approaches, never directly compared across monaural and diotic conditions.

Despite the headway we have made toward delineating markers of musicianship on central auditory processing, the extent to which these enhancements involve diotic sound processing, if at

Received Dec. 11, 2012; revised Sept. 5, 2013; accepted Sept. 11, 2013.

Author contributions: A.P.-C., E.H., and N.K. designed research; A.P.-C. and E.H. performed research; A.P.-C., E.H., and N.K. contributed unpublished reagents/analytic tools; A.P.-C., D.L.S., and E.H. analyzed data; A.P.-C., D.L.S., E.H., and N.K. wrote the paper.

This work was supported by an Undergraduate Academic Research Grant to E.H. and the Knowles Hearing Center and National Science Foundation 0921275 to N.K. We thank the subjects for their participation in this study and Dylan Levy for his assistance with data analysis.

Correspondence should be addressed to Dr. Nina Kraus, Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, IL 60208. E-mail: nkraus@northwestern.edu. www.brainvolts.northwestern.edu.

DOI:10.1523/JNEUROSCI.5700-12.2013

Copyright © 2013 the authors 0270-6474/13/3316741-07\$15.00/0

all, remains unknown. To this aim, we compared musicians' and non-musicians' speech-evoked auditory brainstem responses (ABRs) across monaural and diotic listening conditions. While musician enhancements have previously been documented in ABRs collected in both monaural (Wong et al., 2007; Bidelman and Krishnan, 2010; Bidelman et al., 2011a,b; Strait et al., 2012, 2013a) and diotic conditions (Musacchia et al., 2007; Parbery-Clark et al., 2009a, 2012a,b; Strait et al., 2013a), we selected a stimulus for which there is no musician advantage when presented monaurally (see Results). In light of the importance of neural response timing for both binaural hearing and hearing in noise (Rance et al., 2007; Tzounopoulos and Kraus, 2009; Anderson et al., 2010), we centered our analyses on measures of neural timing (i.e., latency and consistency) in addition to neural response amplitude. We hypothesized that musicians' enhanced auditory processing reflects strengthened diotic processing. Accordingly, we predicted that musicians would demonstrate faster neural timing and increased response magnitudes as measured by the timing and amplitudes of discrete response peaks (i.e., earlier and larger peaks) and increased neural response consistency (i.e., higher between-trial response similarity) in the diotic relative to the monaural condition. We further predicted that these enhancements in diotic processing would relate to musicians' advantages for hearing in noise.

Materials and Methods

Participants

Thirty subjects (mean age 20 ± 2 years, 11 males) were recruited from the Chicago area. All subjects were native English speakers, had normal hearing thresholds (≤ 15 dB HL from 250 to 8000 Hz), and reported no histories of learning or neurological disorders. Participants were screened for normal IQ as measured by the Test of Nonverbal Intelligence (Brown et al., 1997). Subjects were categorized as musicians and non-musicians ($N = 15$ each). Musicians were self-categorized and had consistently practiced an instrument at least three times a week since 7 years of age. Non-musicians had < 3 years of musical training at any point in their lives. See Table 1 for each group's musical experience. Groups were matched for age ($F_{(1,29)} = 0.855$, $p = 0.477$), nonverbal IQ ($F_{(1,29)} = 2.13$, $p = 0.217$), hearing thresholds ($F_{(1,29)} = 1.017$, $p = 0.552$), and sex ($\chi^2_{(1,29)} = 0.741$, $p = 0.389$). Groups had equivalent neural timing to a 100 μ s click stimulus (Wave V: $p > 0.1$) presented at 31.3 Hz and 70 dB sound pressure level (SPL). All subjects gave informed consent before participating in accordance with the Northwestern University Institutional Review Board and were paid for their participation.

Electrophysiology

Stimuli and recording parameters. Auditory brainstem responses were elicited by a 40 ms speech syllable, /da/, at 70 dB SPL under three conditions: monaural right, monaural left, and diotic presentation. The five-formant speech stimulus was synthesized at a sampling rate of 10 kHz using a Klatt-based synthesizer (Klatt, 1980). The stimulus comprised an initial 10 ms onset burst and voiced formant transition between the consonant and the vowel with a fundamental frequency that linearly increased from 103 to 125 Hz. Voicing began at 5 ms. The first formant increased from 220 to 720 Hz. The second and third formants decreased from 1700 and 2580 to 1240 Hz and 2500 Hz, respectively. The fourth and fifth formants were constant at 3600 and 4500 Hz. While the stimulus was short and did not contain a steady-state vowel, it is perceived as a consonant-vowel syllable.

The responses were collected at a 20 kHz sampling rate using NeuroScan Acquire 4.3 recording system (Compumedics) with four Ag-AgCl scalp electrodes in a vertical montage (Cz active, forehead ground, and linked-earlobe reference). Electrodes were coupled to the skin with Ten20 conductive paste (Weaver) and affixed with medical tape. Contact impedance was 2 k Ω or less across all electrodes. Responses were off-line, bandpass filtered from 70 to 2000 Hz, with a 12 dB/octave filter roll-off. Sweeps with activity exceeding ± 35 μ V were considered artifacts and

Table 1. Musical practice history for musicians: years of musical experience, age at which musical training began, and major instruments played are indicated for all musicians

	Years of musical experience	Age onset, years	Instrument
Musicians			
1	15	6	Violin
2	15	5	Violin
3	12	6	Violin
4	20	6	Cello
5	13	5	Flute
6	13	6	Flute
7	14	4	Piano/cello
8	19	6	Piano/French horn
9	23	5	Violin/piano
10	19	5	Violin/piano
11	19	3	Piano
12	15	6	Piano
13	16	5	Piano
14	17	5	Piano
15	13	6	Piano
Mean	16.2	5.1	
Non-musicians			
1	1	16	Guitar
2	3	12	Voice
3	2	10	Piano
4	2	12	Cello
5	2	11	Trumpet
6	1	12	Trumpet
7–15	0	NA	NA
Mean	0.7	12.2	

excluded. The speech stimulus was presented through insert ear phones (ER-3; Etymotic Research) in all three conditions with alternating polarities to limit the inclusion of stimulus artifact and cochlear microphonic (Gorga et al., 1985; Aiken and Picton, 2008; Skoe and Kraus, 2010). For each stimulus polarity in each condition, two subaverages representing 3000 artifact-free responses were averaged to generate an average response comprising 6000 response trials for each condition.

Diotic and monaural conditions were randomized across participants, ruling out contributions of neural fatigue or adaptation to any between-condition effects observed. Participants watched captioned movies of their choice to facilitate a calm yet wakeful recording session.

Data analysis

Timing and response magnitude. To assess neural response timing and magnitude, we identified major response peaks corresponding to the stimulus onset (peaks V and A, occurring at ~ 7 and 8 ms), spectrotemporally dynamic consonant-vowel transition (i.e., frequency following response, or FFR; peaks D–F, occurring at ~ 23 , 31, and 40 ms, respectively) and offset (peak O, occurring at ~ 47 ms; see Fig. 1). Peaks were first identified by two independent peak pickers, after which the first author compared their judgments. The two peak pickers' judgments were identical across all peaks in all subjects except two instances; in these cases, the first author arbitrated. All peak pickers were blind to subjects' groups. All peaks were easily identifiable in all participants in all conditions, with peak minima extending beyond the magnitudes of the inter-peak intervals. Peak minima also extended beyond the noise floor (i.e., magnitude of the prestimulus period).

Response consistency. The consistency of an individual's response to the speech stimulus was measured over the 11–42 ms portion of the recording period—a period encompassing the FFR. Three hundred iterations of randomly selected, subtracted-polarity pairs of 3000 subaverages were created and the degree of correlation (Pearson's r) between the pairs was calculated. These correlations were then averaged to form a composite response consistency score. All processing was performed in MATLAB (MathWorks). Higher correlation coefficients indicate greater neural re-

sponse consistency. All statistical analyses were performed on Fisher transformed z -values.

Speech in noise

Hearing in noise. The Hearing in Noise Test (Biologic Systems; Nilsson et al., 1994) is an adaptive speech-in-noise test that uses Bamford–Kowal–Bench phonetically balanced sentences (Bench et al., 1979) superposed on a speech-shaped noise masker (65 dB SPL). The noise is acoustically fixed, being identical both within and across trials. Participants are instructed to ignore the noise and repeat 20 short semantically and syntactically simple sentences (e.g., “Sugar is very sweet”) that are presented from a loudspeaker placed 1 m directly ahead of the participant. Sentences are counted as correct only when all the words are repeated correctly. The intensity level of the target sentence varies based on the performance of the participant. Performance is assessed by determining the signal-to-noise ratio defined as the difference between the intensity of the target relative to the background noise at which a participant can repeat 50% of the target items correctly. A lower score reflects greater speech-in-noise ability.

Statistical analysis

Statistical analyses were conducted in SPSS 19.0 (SPSS). Two 3 (condition) \times 6 (peak) \times 2 (group) repeated-measures ANOVAs (RMANOVAs) were conducted to assess effects of condition on (1) peak latencies and (2) peak amplitudes in musicians and non-musicians. A 3 (condition) \times 2 (group) RMANOVA was conducted to assess effects of condition on response consistency in musicians and non-musicians. *Post hoc* one-way ANOVAs and paired t tests were conducted to define statistically significant ($p < 0.05$) interaction effects. A one-way ANOVA was used to compare musicians’ to non-musicians’ speech-in-noise perception. Relationships among speech-in-noise perception and neural response characteristics were explored using Pearson’s correlations. Assumptions of normality, linearity, outliers, and multicollinearity were met for all analyses, assessed by normality plots, Shapiro–Wilks test, Mahalanobis distances, and formal Variance Inflation Factor, respectively. All reported statistics reflect two-tailed significance values ($\alpha = 0.05$).

Results

Summary of results

While musicians demonstrated faster neural timing and more consistent auditory brainstem responses to diotically presented sounds, musicians and non-musicians were not distinct in response to the same sound when presented monaurally. Furthermore, musicians demonstrated greater enhancements from monaural to diotic conditions relative to non-musicians, evident in musicians’ faster neural timing, greater response magnitudes, and more consistent responses to diotic relative to monaural stimulation. Musicians’ better speech-in-noise perception correlated with faster neural timing and greater response consistency in the diotic but not the monaural conditions.

Brainstem response

Timing

There was a significant main effect of peak ($F_{(5,24)} > 166,500.58$, $p < 0.0001$), with no main effects of condition or group. There were significant interactions between condition and peak ($F_{(10,19)} = 3.32$, $p < 0.02$) and between condition and group ($F_{(2,27)} = 4.62$, $p < 0.02$) as well as a three-way interaction between condition, peak and group ($F_{(10,19)} = 5.76$, $p < 0.001$). Musicians had faster responses than non-musicians to diotic but not to right- or left-monaural stimulation for all three peaks corresponding to the neural response to the formant transition (Fig. 1 and Tables 2 and 3; diotic peak D: $F_{(1,29)} = 2.38$, $p < 0.001$; peak E: $F_{(1,29)} = 2.00$, $p = 0.06$; peak F: $F_{(1,29)} = 1.83$, $p < 0.02$; monaural transition peaks, all $F < 0.5$, $p > 0.25$). Musicians and non-musicians did not differ in response the speech sound onset or offset in any of

the three conditions (all $F < 0.5$, all $p > 0.1$). *Post hoc* paired t tests indicated that musicians’ responses to two of the transition peaks (i.e., D and E) and the offset peak (i.e., O) in the response to diotically presented stimuli occurred earlier than those to monaurally presented stimuli (diotic vs right presentation, peak D: $t_{(14)} = 3.8$, $p < 0.005$; peak E: $t_{(14)} = 2.6$, $p < 0.02$; peak O: $t_{(14)} = 2.3$, $p < 0.05$; diotic vs left presentation, peak D: $t_{(14)} = 3.6$, $p < 0.005$; peak E: $t_{(14)} = 2.6$, $p < 0.02$; peak O: $t_{(14)} = 2.3$, $p < 0.05$). Non-musicians’ responses did not differ for any peak other than F for the diotic versus right monaural comparison and E for the diotic versus left monaural comparison; in both cases, this peak occurred earlier in the diotic than monaural condition (peak F, diotic vs right presentation: $t_{(14)} = 3.0$, $p < 0.02$; peak E, diotic vs left presentation: $t_{(14)} = 2.8$, $p < 0.02$). Onset peaks V and A diotically compared with monaurally presented stimuli did not differ in either group, nor did right- and left-monaural responses (all $t \leq 2.0$, $p \geq 0.07$).

We further quantified the degree of timing shift invoked by diotic stimulation by subtracting the latencies for peaks D–F and O in the left- and right-monaural conditions from the latencies of these same peaks in the diotic condition. Additionally, we computed a mean of D–F shifts to characterize a global formant-transition shift. Negative values indicate that responses to diotically presented stimuli precede those to monaurally presented stimuli. A one-way ANOVA confirmed greater timing shifts between monaural and diotic conditions in musicians relative to non-musicians (peak D: $F_{(29)} = 16.5$, $p < 0.001$; peak E: $F_{(29)} = 5.7$, $p < 0.05$; peak F: $F_{(29)} = 3.2$, $p = 0.08$; transition composite: $F_{(29)} = 17.0$, $p < 0.001$; peak O: $F_{(29)} = 6.5$, $p < 0.05$). Figure 1D displays the transition composite shift for musicians and non-musicians.

Response magnitude

There were significant main effects of condition and peak (both $F > 20.0$, $p < 0.001$), with all six peaks having greater magnitudes in the diotic relative to either monaural condition and four of the six response peaks being larger with left-ear relative to right-ear stimulation. Furthermore, we observed a three-way interaction between condition, peak, and group ($F_{(10,19)} = 2.4$, $p < 0.05$). *Post hoc* paired-samples t tests indicated that musicians’ had marginally significant enhancements in response magnitudes to diotically relative to right-ear but not left-ear monaurally presented stimuli relative to non-musicians for three of the six response peaks (diotic vs right presentation, peak A: $F_{(1,29)} = 3.7$, $p < 0.06$; peak D: $F_{(1,29)} = 8.9$, $p < 0.01$; peak O: $F_{(1,29)} = 3.9$, $p < 0.06$; diotic vs left presentation, all $F < 0.3$, $p > 0.1$).

Response consistency

Main effects of condition ($F_{(2,27)} = 32.1$, $p < 0.0001$) with response consistency being greater in the diotic rather than the monaural conditions, and group ($F_{(1,28)} = 6.6$, $p < 0.02$), with musicians demonstrating greater neural response consistency ($r = 0.24$, $SD = 0.132$) across conditions relative to non-musicians ($r = 0.16$, $SD = 0.116$), were found. There was also a significant group \times condition interaction ($F_{(2,27)} = 13.5$, $p < 0.001$). *Post hoc* one-way ANOVAs indicated that musicians’ greater response consistency across conditions was driven by an enhancement in the diotic condition; musicians had greater neural response consistency than non-musicians only in response to diotic but not monaural stimulation (Fig. 2; diotic: $F_{(1,29)} = 33.4$, $p < 0.001$; right and left: both $F < 0.5$, $p > 0.5$).

Relationships between neural measures and speech-in-noise perception

Musicians demonstrated superior hearing-in-noise ability compared with age-matched non-musicians ($F_{(1,29)} = 6.606, p < 0.01$). Across all subjects, earlier peak timing and greater neural response consistency in response to diotically presented stimuli related to better speech-in-noise perception (peak E timing: $r = 0.46, p < 0.02$; response consistency: $r = -0.591, p < 0.001$). These relationships were exclusive to the diotic condition (right- and left-monaural conditions: all $p > 0.1$).

Discussion

We herein demonstrate musicians' enhancements for processing diotic sounds, with musicians demonstrating faster, larger, and more consistent responses to speech sounds presented simultaneously to both ears but not to the same sounds presented alone to the right or left ears. Furthermore, these physiological indices relate to speech-in-noise perception, suggesting that musicians' advantages may be driven, at least in part, by better processing of diotically presented sounds. Musicians' enhancements to the diotically presented stimulus may reflect music listening's persistent reliance on binaural sound processing; localization of a sound source in ensemble playing or conducting, for example, requires the precise and robust encoding of sounds by both ears to differentiate sounds presented closer to one ear from those presented closer to the other. This extensive auditory experience may equip musicians to preferentially process diotically presented sounds even in the absence of interaural differences providing timing or level cues.

Binaural processing and its malleability with sensory experience

Acoustic signals that arrive at the two ears are initially transmitted ipsilaterally along the auditory pathway. While sound input to each ear at first remains separate, it quickly reaches three points of convergence: the superior olive, the nuclei of the lateral lemniscus, and the inferior colliculus (Bocca, 1955; for review, see Moore, 1991; Wallace et al., 1996; McAlpine et al., 2000). At these points, the auditory system integrates information, detecting differences in the phase, intensity, and timing of the signals from the two ears (Moore, 1991). While cortical auditory structures receive integrated sound input from both ears, their respective weightings within auditory cortex are not predetermined—they can be modulated by auditory experiences early in life (Hogan and Moore, 2003) and the engagement of cognitive functions during the listening process (Shinn-Cunningham et al., 2005).

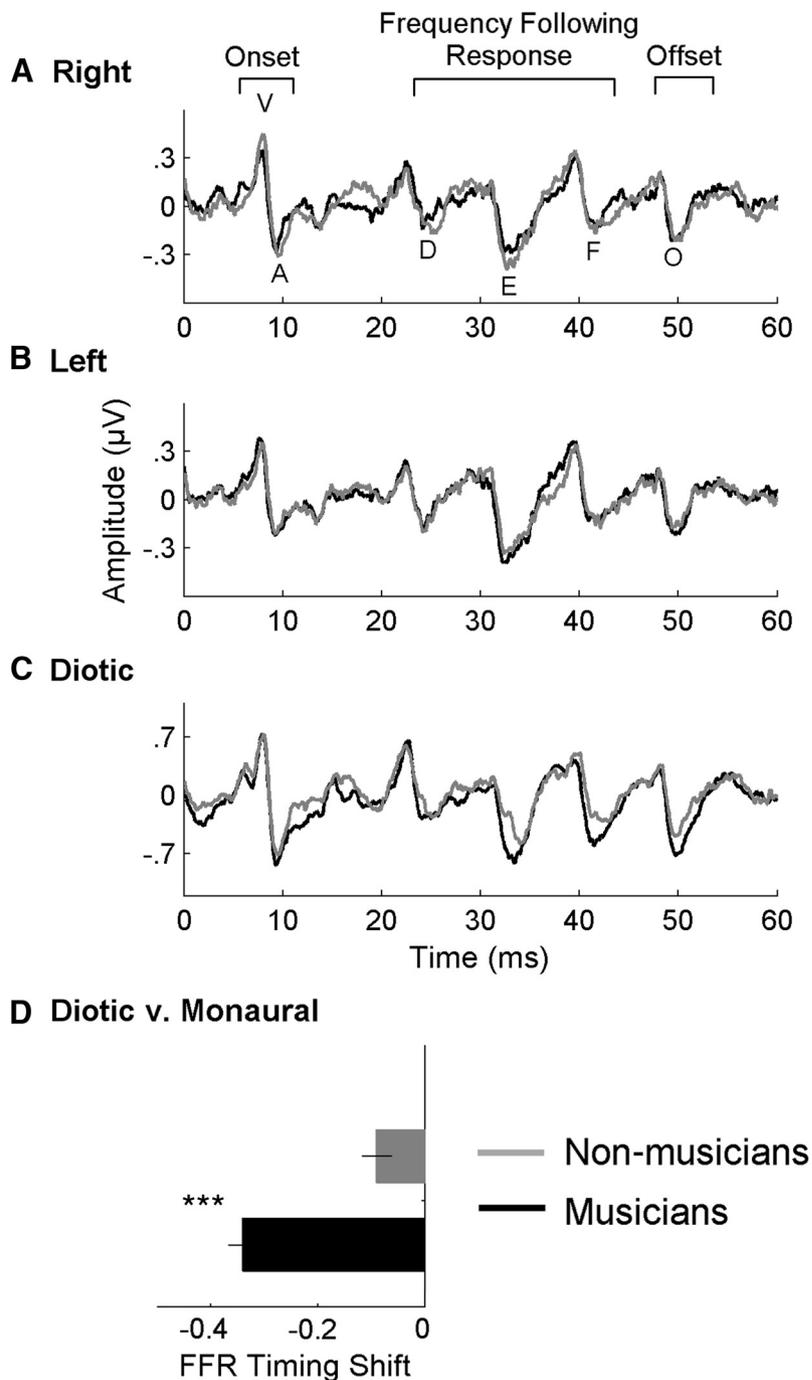


Figure 1. ABR to the stimulus “da.” The speech-evoked ABR for musicians (black) and non-musicians (gray) for right, left, and diotic presentation. *A, B*, Musicians and non-musicians had equivalent peak timing in both monaural paradigms. *C*, Musicians had earlier diotic peak timing than non-musicians in the FFR region of the response. *D*, Musicians have a greater average difference in formant transition peak timing than non-musicians between responses to diotically and monaurally presented stimuli (i.e., average monaural-to-diotic timing differences for peaks D–F), with earlier neural timing in the diotic condition ($F_{(29)} = 17.0, p < 0.001$). Error bars represent ± 1 standard error. *** $p < 0.001$.

While the diotic stimulus presentation used here does not tap into traditional binaural processing cues such as timing and level differences, it may be that musicians' diotic advantage reflects binaural processing's known experience-related malleability. Although the structural development of the human auditory brainstem is thought to be complete within the first two years of life (Moore et al., 1995; Moore and Linthicum, 2007) and is guided by genetics (Clopton and Silverman, 1977; Taniguchi, 1981), how

Table 2. Diotic presentation group differences: group means (SDs) and significance values for measures of neural timing (ms) and response consistency (r-value) in the diotic condition

	Mean (SD) Musicians	Mean (SD) Non-musicians	Group p value
Timing: V	7.02 (0.21)	7.07 (0.18)	0.982
A	8.28 (0.27)	8.30 (0.29)	0.718
D	23.07 (0.43)	23.09 (0.38)	<0.001
E	31.69 (0.65)	31.72 (0.84)	0.028
F	40.43 (0.51)	40.74 (0.69)	0.005
Offset	48.51 (0.33)	48.51 (0.31)	0.075
Response consistency	0.413 (0.13)	0.237 (0.08)	<0.001

Table 3. Monaural presentation group differences: group means (SDs) and group difference p values for timing (ms) and response consistency (r-value) in right and left monaural conditions

	Mean (SD) Musicians Left/right	Mean (SD) Non-musicians Left/right	Group p value Left/right
Timing: V	7.07 (0.20)/7.08 (0.19)	7.12 (0.29)/7.10 (0.16)	0.479/0.981
A	8.27 (0.21)/8.28 (0.41)	8.31 (0.26)/8.32 (0.33)	0.747/0.790
D	23.10 (0.33)/23.19 (0.22)	23.13 (0.41)/23.15 (0.27)	0.154/0.105
E	31.75 (0.47)/31.80 (0.48)	31.76 (0.51)/31.82 (0.51)	0.293/0.548
F	40.61 (0.50)/40.54 (0.37)	40.79 (0.48)/40.82 (0.64)	0.145/0.398
Offset	48.60 (0.52)/48.59 (0.36)	48.73 (0.71)/48.77 (0.44)	0.745/0.220
Response consistency	0.216 (0.22)/0.218 (0.19)	0.200 (0.11)/0.209 (0.12)	0.729/0.590

There were no group differences in any neural measure in the monaural conditions.

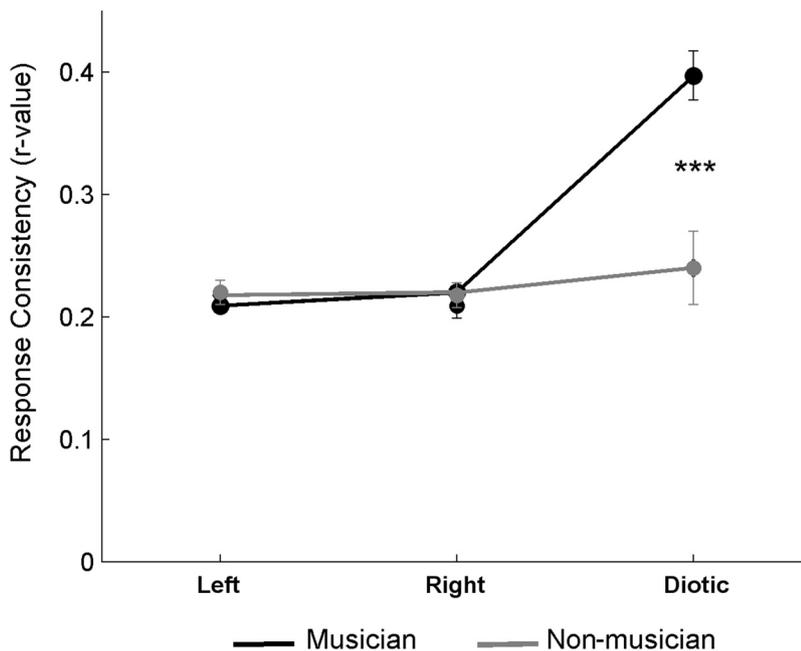


Figure 2. Response consistency measured across listening conditions. While musicians (black) and non-musicians (gray) had equally consistent responses in both monaural conditions, musicians have more consistent responses in the diotic condition. Error bars represent ± 1 standard error. $***p < 0.001$.

the auditory system makes use of this circuitry to achieve sound identification and localization is experience dependent. The functional organization of brainstem nuclei involved in binaural hearing is modulated by sensory experiences that occur during development (Clopton and Silverman, 1977; Knudsen et al., 1984a, b; Knudsen, 1985; for review, see King et al., 2000) and, to a lesser extent, adulthood (Kacelnik et al., 2006). Because of this, even the mature auditory system is capable of adapting to changes

in the balance between the two ears (Bauer et al., 1966; Florentine, 1976; Hofman et al., 1998; Shinn-Cunningham et al., 1998; Van Wanrooij and Van Opstal, 2005). Experience-related changes in subcortical binaural hearing structures may stem, at least in part, from the structural and functional organization of auditory cortex that continues into adolescence (Gleeson and Walsh, 2000; Moore and Guan, 2001; Moore and Linthicum, 2007) via top-down modulation of neuronal response properties. In fact, deactivating descending inputs to inferior colliculus prevents sound localization using binaural hearing cues, evident in conditions where the balance in hearing between the two ears has been experimentally altered (Bajo et al., 2010).

While we have interpreted our results in the context of training-related changes in musicians' binaural sound processing, group comparisons cannot disentangle innate and training-related factors contributing to musicians' binaural processing enhancements. Future studies, most notably longitudinal work, should define both the developmental trajectory of monaural and binaural auditory processing advantages in musicians and their direct relationships to training over and above innate predispositions. Additionally, it is possible that our stimulus presentation mode—lacking interaural timing or level cues—enabled musicians to take advantage of their “better ear.” This alternative interpretation would support the possibility that, when presented with diotic stimulation, musicians preferentially benefited from the use of their dominant ear, resulting in an enhanced physiological response.

Interpretation according to previous monaural and binaural approaches in musicians

While these results evidence enhanced diotic but not monaural processing in musicians, musician enhancements have previously been documented in ABRs collected in both monaural (Wong et al., 2007; Strait et al., 2012, 2013a) and diotic conditions (Musacchia et al., 2007; Parbery-Clark et al., 2009a, 2012a,b; Strait et al., 2013a). To test our hypothesis that musicians' enhanced speech-in-noise perception reflects strengthened processing to diotically presented sounds, the present study intentionally used a stimulus that did not elicit a musician advantage when presented monaurally. This stimulus was unique in that it did not contain the acoustically rich vowel portion of the syllable, which is known to elicit more robust neural encoding of the spectral components of speech in musicians (Musacchia et al., 2007; Wong et al., 2007; Parbery-Clark et al., 2009a, 2012a; Strait et al., 2012; Strait and Kraus, 2013). The absence of this vowel region could account for the lack of monaural enhancement observed in musicians to this stimulus due to decreased acoustic content and less backward masking of the formant transition by the broadband vowel. Still, further work should disentangle acoustic parameters that successfully induce musician enhancements in both monaural and diotic conditions from those that do not. Outcomes may shed light on musical training as a model of experience-related neuroplasticity, specifying the markers of musicianship against aspects of auditory processing that remain unaffected.

Practical applications

These results stimulate further investigation into how auditory training might be used in clinical populations that demonstrate compromised binaural processing. Binaural hearing's relationship to speech-in-noise perception (Nábělek and Robinson, 1982; Ellermeier and Hellbrück, 1998; Strouse et al., 1998; Hawley et al., 2004) encourages its relevance for school-aged children given that their everyday learning occurs in noisy classroom environments; in fact, a typical in-session elementary school classroom is ~60 dB (Rosenberg et al., 1999; Bradley, 2005; Cameron et al., 2006). Given that conversational voice levels are ~50 dB, it is not surprising that low signal-to-noise ratios can make hearing what the teacher is saying a challenging task (Barton, 1989; Blair, 1990). Our results suggest that strengthening the neural pathways involved in binaural processing through auditory training, such as music lessons, may alleviate some of the classroom difficulties faced by children that involve understanding speech in noise.

Future directions

Future work aimed at defining additional behavioral and neural assessments of binaural processing differences in musicians and non-musicians, including assessments comprising more traditional interaural timing and level difference paradigms, will expand our understanding of the extent to which binaural processing is experience dependent. Furthermore, this work might consider subgroups of musicians, differentiated by the degree to which they depend on binaural hearing cues in their musical practice. While musicians have enhanced sound localization relative to non-musicians (Tervaniemi et al., 2006), there is also the possibility of between-musician group differences. Conductors, for example, must regularly locate sound sources within a busy orchestra and demonstrate strengthened sound localization (Münte et al., 2001); accordingly, conductors may demonstrate binaural processing enhancements above and beyond those reported here in performing musicians. Similarly, orchestral musicians (e.g., string players) may demonstrate enhancements relative to soloists (e.g., pianists). Throughout this work, musicians' cognitive and speech-in-noise perceptual benefits might be assessed using both binaural and monaural approaches, in addition to assessing binaural unmasking and sound localization, which are both dependent on hearing with both ears.

Notes

Supplemental information can be found online at <http://www.soc.northwestern.edu/brainvolts/>.

References

- Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245:35–47. [CrossRef Medline](#)
- Anderson S, Skoe E, Chandrasekaran B, Kraus N (2010) Neural timing is linked to speech perception in noise. *J Neurosci* 30:4922–4926. [CrossRef Medline](#)
- Bajo VM, Nodal FR, Moore DR, King AJ (2010) The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 13:253–260. [Medline](#)
- Barton L (1989) Sound levels in occupied elementary school classrooms. Unpublished master's thesis, Utah State University.
- Bauer RW, Matuzsa JL, Blackmer F, Glucksberg S (1966) Noise localization after unilateral attenuation. *J Acoust Soc Am* 40:441–444. [CrossRef](#)
- Bench J, Kowal A, Bamford J (1979) The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br J Audiol* 13:108–112. [CrossRef Medline](#)
- Bidelman GM, Krishnan A (2010) Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res* 1355:112–125. [CrossRef Medline](#)
- Bidelman GM, Gandour JT, Krishnan A (2011a) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 23:425–434. [CrossRef Medline](#)
- Bidelman GM, Gandour JT, Krishnan A (2011b) Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn* 77:1–10. [CrossRef Medline](#)
- Blair J (1990) Front-row seating is not enough for classroom listening. How the student with hearing loss can succeed in college: a handbook for students, families, and professionals Washington, DC, Alexander Graham Bell Association for the Deaf.
- Bocca E (1955) Binaural hearing: another approach. *Laryngoscope* 65:1164–1171. [Medline](#)
- Bradley JS (2005) Does the classroom assist or impede the learning process? *Can Assoc Principals J* 13:32–34.
- Brown L, Sherbenou RJ, Johnsen SK (1997) Test of nonverbal intelligence-3 (TONI-3). Austin, TX: Pro-Ed.
- Cameron S, Dillon H, Newall P (2006) The listening in spatialized noise test: an auditory processing disorder study. *J Am Acad Audiol* 17:306–320. [CrossRef Medline](#)
- Clopton BM, Silverman MS (1977) Plasticity of binaural interaction: II. Critical period and changes in midline response. *J Neurophysiol* 40:1275–1280. [Medline](#)
- Davis AC, Haggard MP (1982) Some implications of audiological measures in the population for binaural aiding strategies. *Scand Audiol Suppl* 15:167–179. [Medline](#)
- Davis A, Haggard M, Bell I (1990) Magnitude of diotic summation in speech-in-noise tasks: performance region and appropriate baseline. *Br J Audiol* 24:11–16. [CrossRef Medline](#)
- Ellermeier W, Hellbrück J (1998) Is level irrelevant in “irrelevant speech”? Effects of loudness, signal-to-noise ratio, and binaural unmasking. *J Exp Psychol Hum Percept Perform* 24:1406–1414. [CrossRef Medline](#)
- Florentine M (1976) Relation between lateralization and loudness in asymmetrical hearing losses. *J Am Audiol Soc* 1:243–251. [Medline](#)
- Fujiki N, Jousmaki V, Hari R (2002) Neuromagnetic responses to frequency-tagged sounds: a new method to follow inputs from each ear to the human auditory cortex during binaural hearing. *J Neurosci* 22:RC205. [Medline](#)
- Gleeson JG, Walsh CA (2000) Neuronal migration disorders: from genetic diseases to developmental mechanisms. *Trends Neurosci* 23:352–359. [CrossRef Medline](#)
- Gorga M, Abbas P, Worthington D (1985) Stimulus calibration in ABR measurements. In: *The auditory brainstem response* (Jacobsen J, ed), pp 49–62. San Diego: College Hill Press.
- Grothe B (2003) New roles for synaptic inhibition in sound localization. *Nat Rev Neurosci* 4:540–550. [CrossRef Medline](#)
- Hanna-Pladdy B, MacKay A (2011) The relation between instrumental musical activity and cognitive aging. *Neuropsychology* 25:378–386. [CrossRef Medline](#)
- Hawley ML, Litovsky RY, Culling JF (2004) The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J Acoust Soc Am* 115:833–843. [CrossRef Medline](#)
- Hofman PM, Van Riswick JG, Van Opstal AJ (1998) Relearning sound localization with new ears. *Nat Neurosci* 1:417–421. [CrossRef Medline](#)
- Hogan SC, Moore DR (2003) Impaired binaural hearing in children produced by a threshold level of middle ear disease. *J Assoc Res Otolaryngol* 4:123–129. [CrossRef Medline](#)
- Hudspeth AJ (1997) How hearing happens minireview. *Neuron* 19:947–950. [CrossRef Medline](#)
- Kacelnik O, Nodal FR, Parsons CH, King AJ (2006) Training-induced plasticity of auditory localization in adult mammals. *PLoS Biol* 4:e71. [CrossRef Medline](#)
- Kaplan H, Pickett JM (1981) Effects of dichotic/diotic versus monotic presentation on speech understanding in noise in elderly hearing-impaired listeners. *Ear Hear* 2:202–207. [CrossRef Medline](#)
- King AJ, Parsons CH, Moore DR (2000) Plasticity in the neural coding of auditory space in the mammalian brain. *Proc Natl Acad Sci U S A* 97:11821–11828. [CrossRef Medline](#)
- Klatt DH (1980) Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 67:971–995. [CrossRef](#)
- Knudsen EI (1985) Experience alters the spatial tuning of auditory units in the optic tectum during a sensitive period in the barn owl. *J Neurosci* 5:3094–3109. [Medline](#)
- Knudsen EI, Knudsen PF, Esterly SD (1984a) A critical period for the recovery

- ery of sound localization accuracy following monaural occlusion in the barn owl. *J Neurosci* 4:1012–1020. [Medline](#)
- Knudsen EI, Esterly SD, Knudsen PF (1984b) Monaural occlusion alters sound localization during a sensitive period in the barn owl. *J Neurosci* 4:1001–1011. [Medline](#)
- Kraus N, Chandrasekaran B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605. [CrossRef Medline](#)
- Kühnis J, Elmer S, Meyer M, Jäncke L (2013) Musicianship boosts perceptual learning of pseudoword-chimeras: an electrophysiological approach. *Brain Topogr* 26:110–125. [CrossRef Medline](#)
- McAlpine D, Jiang D, Shackleton TM, Palmer AR (2000) Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. *J Neurophysiol* 83:1356–1365. [Medline](#)
- Moore DR (1991) Anatomy and physiology of binaural hearing. *Int J Audiol* 30:125–134. [CrossRef](#)
- Moore JK, Guan YL (2001) Cytoarchitectural and axonal maturation in human auditory cortex. *J Assoc Res Otolaryngol* 2:297–311. [CrossRef Medline](#)
- Moore JK, Linthicum FH Jr (2007) The human auditory system: a timeline of development. *Int J Audiol* 46:460–478. [CrossRef Medline](#)
- Moore JK, Perazzo LM, Braun A (1995) Time course of axonal myelination in the human brainstem auditory pathway. *Hear Res* 87:21–31. [CrossRef Medline](#)
- Münste TF, Kohlmetz C, Nager W, Altenmüller E (2001) Superior auditory spatial tuning in conductors. *Nature* 409:580. [CrossRef Medline](#)
- Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci U S A* 104:15894–15909. [CrossRef Medline](#)
- Nábělek AK, Robinson PK (1982) Monaural and binaural speech perception in reverberation for listeners of various ages. *J Acoust Soc Am* 71:1242–1248. [CrossRef Medline](#)
- Nilsson M, Soli SD, Sullivan JA (1994) Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95:1085–1099. [CrossRef Medline](#)
- Parbery-Clark A, Skoe E, Kraus N (2009a) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci* 29:14100–14107. [CrossRef Medline](#)
- Parbery-Clark A, Skoe E, Lam C, Kraus N (2009b) Musician enhancement for speech-in-noise. *Ear Hear* 30:653–661. [CrossRef Medline](#)
- Parbery-Clark A, Strait DL, Anderson S, Hittner E, Kraus N (2011) Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. [CrossRef Medline](#)
- Parbery-Clark A, Anderson S, Hittner E, Kraus N (2012a) Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front Aging Neurosci* 4:30. [Medline](#)
- Parbery-Clark A, Tierney A, Strait DL, Kraus N (2012b) Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219:111–119. [CrossRef Medline](#)
- Plomp R, Mimpen AM (1979) Improving the reliability of testing the speech reception threshold for sentences. *Int J Audiol* 18:43–52. [CrossRef](#)
- Rance G, Barker E, Mok M, Dowell R, Rincon A, Garratt R (2007) Speech perception in noise for children with auditory neuropathy/dys-synchrony type hearing loss. *Ear Hear* 28:351–360. [CrossRef Medline](#)
- Rosenberg GG, Blake-Rahter P, Heavner J, Allen L, Redmond BM, Phillips J, Stigers K (1999) Improving classroom acoustics (ICA): a three-year FM sound field classroom amplification study. *J Educ Audiol* 7:22.
- Shinn-Cunningham BG, Durlach NI, Held RM (1998) Adapting to super-normal auditory localization cues. I. Bias and resolution. *J Acoust Soc Am* 103:3656–3666. [CrossRef Medline](#)
- Shinn-Cunningham BG, Ihlefeld A, Satyavarta LE (2005) Bottom-up and top-down influences on spatial unmasking. *Acta Acustica* 91:967–979.
- Shook A, Marian V, Bartolotti J, Schroeder SR (2013) Musical experience influences statistical learning of a novel language. *Am J Psychol* 126:95–104. [CrossRef Medline](#)
- Skoe E, Kraus N (2010) Auditory brainstem response to complex sounds: a tutorial. *Ear Hear* 31:302–324. [CrossRef Medline](#)
- Strait DL, Kraus N (2013) Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hear Res pii: S0378-5955(13)00195-0*. [CrossRef Medline](#)
- Strait DL, Kraus N, Parbery-Clark A, Ashley R (2010) Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res* 261:22–29. [CrossRef Medline](#)
- Strait DL, O'Connell S, Parbery-Clark A, Kraus N (2013a) Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb Cortex*. Advance online publication. Retrieved April 18, 2013. doi:10.1093/cercor/bht103.
- Strait DL, O'Connell S, Parbery-Clark A, O'Connell S, Kraus N (2013b) Biological impact of preschool music classes on processing speech in noise. *Dev Cogn Neurosci* 6C:51–60. [Medline](#)
- Strait DL, Parbery-Clark A, Hittner E, Kraus N (2012) Musical training during early childhood enhances the neural encoding of speech in noise. *Brain Lang* 123:191–201. [CrossRef Medline](#)
- Strouse A, Ashmead DH, Ohde RN, Grantham DW (1998) Temporal processing in the aging auditory system. *J Acoust Soc Am* 104:2385–2399. [CrossRef Medline](#)
- Taniguchi I (1981) Plastic changes in inferior colliculus following cochlear destruction. In: *Neuronal mechanisms of hearing* (Syka J, Aitkin L, eds), pp 377–380. New York: Plenum.
- Tervaniemi M, Castaneda A, Knoll M, Uther M (2006) Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. *Neuroreport* 17:1225–1228. [CrossRef Medline](#)
- Tzounopoulos T, Kraus N (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62:463–469. [CrossRef Medline](#)
- Van Wanrooij MM, Van Opstal AJ (2005) Relearning sound localization with a new ear. *J Neurosci* 25:5413–5424. [CrossRef Medline](#)
- Wallace MT, Wilkinson L, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. *J Neurophysiol* 76:1246–1266. [Medline](#)
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422. [Medline](#)
- Zendel BR, Alain C (2009) Concurrent sound segregation is enhanced in musicians. *J Cogn Neurosci* 21:1488–1498. [CrossRef Medline](#)