

# Musicians' Enhanced Neural Differentiation of Speech Sounds Arises Early in Life: Developmental Evidence from Ages 3 to 30

Dana L. Strait<sup>1,2</sup>, Samantha O'Connell<sup>2</sup>, Alexandra Parbery-Clark<sup>2,3</sup> and Nina Kraus<sup>1,2,3,4,5</sup>

<sup>1</sup>Institute for Neuroscience, <sup>2</sup>Auditory Neuroscience Laboratory, <sup>3</sup>Department of Communication Sciences, <sup>4</sup>Department of Neurobiology and Physiology and <sup>5</sup>Department of Otolaryngology, Northwestern University, Evanston, IL, USA

Address correspondence to Nina Kraus, Auditory Neuroscience Laboratory, 2240 Campus Drive, Evanston, IL 60208, USA.  
Email: nkraus@northwestern.edu; <http://www.brainvolts.northwestern.edu>

**The perception and neural representation of acoustically similar speech sounds underlie language development. Music training hones the perception of minute acoustic differences that distinguish sounds; this training may generalize to speech processing given that adult musicians have enhanced neural differentiation of similar speech syllables compared with nonmusicians. Here, we asked whether this neural advantage in musicians is present early in life by assessing musically trained and untrained children as young as age 3. We assessed auditory brainstem responses to the speech syllables /ba/ and /ga/ as well as auditory and visual cognitive abilities in musicians and nonmusicians across 3 developmental time-points: preschoolers, school-aged children, and adults. Cross-phase analyses objectively measured the degree to which subcortical responses differed to these speech syllables in musicians and nonmusicians for each age group. Results reveal that musicians exhibit enhanced neural differentiation of stop consonants early in life and with as little as a few years of training. Furthermore, the extent of subcortical stop consonant distinction correlates with auditory-specific cognitive abilities (i.e., auditory working memory and attention). Results are interpreted according to a corticofugal framework for auditory learning in which subcortical processing enhancements are engendered by strengthened cognitive control over auditory function in musicians.**

**Keywords:** attention, auditory, brain, language, musical training

## Introduction

Experience-dependent characteristics of musicians' speech-evoked auditory brainstem function have been of significant interest in recent years (for review, see Kraus and Chandrasekaran 2010; Patel 2011; Strait and Kraus 2011). Accumulating evidence substantiates musicians' faster, more robust and more precise auditory brainstem responses (ABRs) to speech and other communication sounds (Musacchia et al. 2007; Wong et al. 2007; Bidelman et al. 2009; Strait et al. 2009; Parbery-Clark, Anderson, et al. 2012a), more distinct responses to similar speech sounds (Parbery-Clark et al. 2011; Parbery-Clark, Tierney, et al. 2012), and increased neural resilience in degraded listening environments (Parbery-Clark et al. 2009; Bidelman and Krishnan 2010; Strait et al. 2012). Although we have made considerable headway toward defining aspects of subcortical auditory function that are enhanced in musicians, efforts have primarily focused on comparisons between adult musicians and nonmusicians. We have not sufficiently addressed whether these enhancements are observable early in life, when musical expertise and its underlying biology are under rapid development. Subcortical auditory processing advantages in child musicians could strengthen aspects of classroom learning given the prerequisites for

learning in language-dominated classroom environments and the importance of accurate subcortical speech-sound encoding for the development of language-related skills (e.g., reading) (Banai et al. 2009; Hornickel et al. 2009, 2011; Billiet and Bellis 2011; Strait et al. 2011; Rocha-Muniz et al. 2012).

We recently reported enhanced subcortical distinction of similar speech sounds (e.g., /ba/ and /ga/) in adult musicians (Parbery-Clark, Tierney, et al. 2012). The distinct perception and neural representation of similar speech sounds are key components in the development of language abilities (Bradley and Bryant 1983; Kraus et al. 1995, 1996; Siok and Fletcher 2001; Hornickel et al. 2009; Chobert, Francois, Habib, et al. 2012). We proposed that musicians' strengthened neural speech-sound distinction has the potential to benefit the acquisition of language-related skills, such as reading and speech-in-noise perception, during development. If this is the case, the impact of musical training on speech-sound processing should be quantifiable in school-aged children.

Here, we addressed whether musicians' enhanced subcortical differentiation of acoustically similar speech sounds is present during childhood by assessing auditory brainstem function in musicians and nonmusicians corresponding to 3 developmental time-points: preschoolers (ages 3–5 years), school-aged children (ages 7–13 years), and young adults (ages 18–30 years). We predicted that musicians have more precise subcortical encoding of speech syllables early in life, apparent in musician children with just a few years of training. We further assessed cognitive abilities (i.e., working memory and attention) in both auditory and visual domains. We predicted that the extent of neural speech-sound processing enhancement observed in musicians would relate to their auditory-specific cognitive performance. This prediction reflects our argument that musicians' sensory processing enhancements stem, at least in part, from greater cognitive control over auditory function by means of a strengthened efferent auditory system (Kraus and Chandrasekaran 2010; Strait et al. 2010; Strait and Kraus 2011).

## Methods

### Participants

All experimental procedures were approved by the Northwestern University Institutional Review Board. Seventy-six normal hearing children and adults (<20 dB HL pure tone thresholds at octave frequencies from 125 to 8000 Hz) between the ages of 3–30 years participated in this study and were grouped by age: preschoolers (3–5 years old,  $N=21$ ), school-aged children (7–13 years old,  $N=26$ ), and adults (18–30 years old,  $N=29$ ). Although all adults were tested on the neurophysiological paradigm, only 17 of the 29 adults (10 musicians) received the full cognitive assessment. Participants and, in the

case of minors, legal guardians provided informed consent and assent. Participants were monetarily compensated for their time. No participant reported a history of neurological or learning abnormalities. Inclusionary criteria also included normal wave V click-evoked ABR latencies and normal IQ (see Cognitive Testing section).

Subjects within each age group were further categorized as musicians (Mus) or nonmusicians (NonMus). Musicians for each age group were self-categorized and were currently undergoing private or, in the case of preschoolers, group music training (e.g., Kindermusik, Orff music classes). Adult musicians ( $N=14$ ) began music training by or before age 7 ( $M=5.4$  years,  $SD=0.75$ ; years practiced  $M=16.7$ ,  $SD=3.50$ ) and had no significant lapses in their practice histories. School-aged musicians ( $N=13$ ) began music training by or before age 6 ( $M=4.9$  years,  $SD=0.81$ ) and had consistently practiced for a minimum of 3 years ( $M=7.2$  years,  $SD=2.43$ ). Preschool musicians ( $N=12$ ) had consistently practiced for a minimum of 12 consecutive months leading up to the date of test ( $M=3.3$  years,  $SD=1.16$ ). Although the style of training varied across preschoolers (e.g., group or private lessons, training on one or many instruments, focus on tonal or percussive instruments, incorporation of singing), all musicians received weekly instruction (on average 1.12 h per week) and used at-home practice materials at least 4 days per week. Preschool nonmusicians had no music training during the year leading up to the test and  $\leq 6$  months over the course of their lives. In fact, only 2 preschool nonmusicians had any degree of music training (group music classes for 3 and 6 months, respectively). School-aged child and adult nonmusicians were self-categorized; nonmusicians had  $< 3$  years of accumulated musical experience throughout their lifespans (children:  $M=0.3$ ,  $SD=0.85$ ; adults:  $M=0.5$ ,  $SD=0.92$ ). Adult musicians practiced a minimum of 3 days per week for  $\geq 1$  h per session, whereas school-aged child musicians practiced for a minimum of 20 min per day, 5 days per week.

Within each age group, Mus and NonMus did not differ according to age (preschoolers:  $F_{1,20}=2.2$ ,  $P=0.15$ ; children:  $F_{1,25}=0.4$ ,  $P=0.54$ ; adults:  $F_{1,28}=0.4$ ,  $P=0.54$ ), sex (adults:  $\chi^2=2.0$ ,  $P=0.25$ ; children:  $\chi^2=2.5$ ,  $P=0.24$ ; preschoolers:  $\chi^2=1.3$ ,  $P=0.39$ ), or IQ (preschoolers:  $F_{1,20}=0.3$ ,  $P=0.59$ ; children:  $F_{1,25}=0.3$ ,  $P=0.57$ ; adults:  $F_{1,28}=1.1$ ,  $P=0.30$ ). School-aged and preschool Mus and NonMus did not differ by socioeconomic status as inferred by maternal education (preschoolers:  $F_{1,20}=0.3$ ,  $P=0.61$ ; children:  $F_{1,25}=1.5$ ,  $P=0.24$ ; see Stevens et al. (2009) for discussion regarding the predictive value of maternal education for inferring a child's socioeconomic status).

### Cognitive Testing

#### IQ

IQ was assessed in preschoolers using the *Peabody Picture Vocabulary Test* (Dunn and Dunn 1997) and in school-aged children using the 2-scale IQ test comprising verbal and nonverbal subtests in the *Wechsler Abbreviated Scale of Intelligence* (Wechsler 1999). IQ was assessed in adults using the *Test of Nonverbal Intelligence* (Brown et al. 1997).

#### Working Memory

We tested auditory working memory (AWM) in school-aged children and adults using the AWM subtest of the *Woodcock-Johnson III Test of Cognitive Abilities* (Woodcock et al. 2001). Participants reordered a dictated series of intermixed digits and nouns by first repeating the nouns and then the digits in their respective sequential orders (e.g., the correct ordering of the following sequence, “4, salt, fox, 7, stove, 2” is “salt, fox, stove, 4, 7, 2”), with a maximum difficulty of 8 items (4 nouns and 4 digits). Age-normed standard scores were used for all statistical analyses.

We tested visual working memory (VWM) in school-aged children and a subset of the adults using the visual span subtest of the *Colorado Assessment Tests 1.2* (Davis and Keller 2002). Participants were instructed to monitor a computer screen displaying 8 blue boxes that sequentially changed color. Participants were asked to click on the boxes in the order in which they changed color. The number of

boxes changing color increased with successive correct replies. Although participants completed both forward and reversed conditions, the reversed condition is represented here as VWM because, like AWM, it requires the manipulation of stored input.

#### Attention

Auditory and visual attention tasks (AAtt and VAtt, respectively) were assessed in school-aged children and a subset of the adults using the *IHR Multicentre Battery for Auditory Processing's* attention subtests (Barry et al. 2010). The tests were administered in a sound-attenuated booth using a laptop computer that was placed 60 cm in front of the participant. Responses were recorded using a 3-button response box. Stimuli were presented diotically through Sennheiser HD 25-1 headphones at 70 dB SPL and were accompanied by animated visual stimuli. For the auditory attention task (AAtt), participants were instructed to watch the computer screen that displayed a cartoon character. Knowing that the visual scene would not change, they were asked to listen for a “beep” and press the center button on the response box with their dominant hand as soon as they heard it. The beep was considered the target stimulus. Participants were cued by a second auditory stimulus (a “siren”) on some trials, which always preceded the target stimulus; participants were asked not to respond to that cue. For the visual attention task (VAtt), participants were instructed to monitor the character for movement and press the center button on the response box with their dominant hand as soon as they saw the character raise its arms. The arm raise was considered the target stimulus. Participants were cued by a second visual stimulus (the changing of the character's shirt color) on some trials, which always preceded the target stimulus; participants were asked not to respond to that cue. Reaction times for both VAtt and AAtt tasks were measured in milliseconds.

### Auditory Brainstem Response Recordings

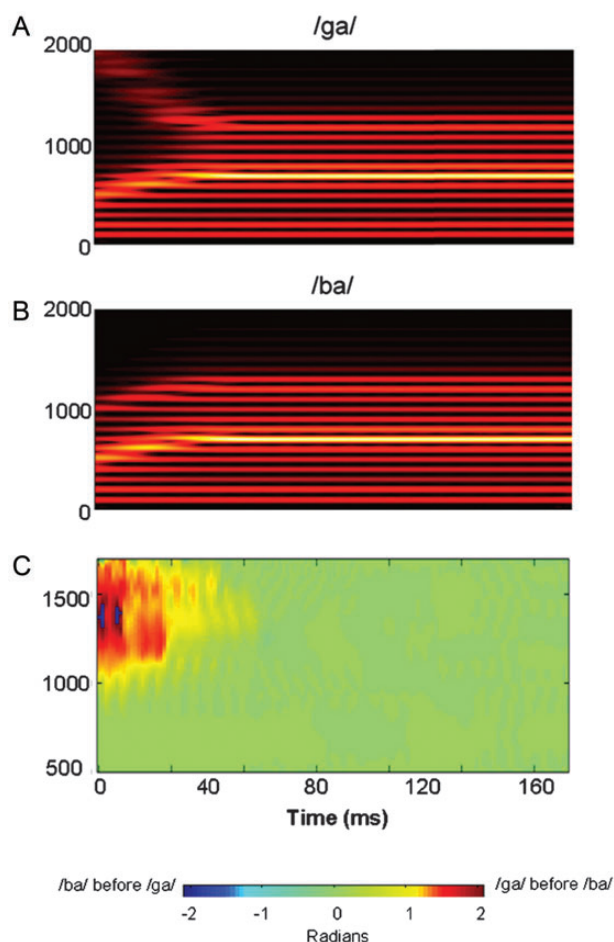
#### Stimuli

The speech syllables /ga/ and /ba/ were constructed using a Klatt-based synthesizer (Klatt 1980). Each syllable is 170 ms in duration with an unchanging fundamental frequency ( $F_0=100$  Hz). For the first 50 ms of both syllables (i.e., for the transition between the consonant stop burst and the vowel), the first and third harmonics change over time ( $F_1=400$ – $720$  Hz;  $F_3=2580$ – $2500$  Hz) while the fourth, fifth, and sixth harmonics remain steady ( $F_4=3300$  Hz;  $F_5=3750$  Hz;  $F_6=4900$  Hz). The syllables are distinguished by the trajectory of their second formants: /ga/ falls from 2480 to 1240 Hz while /ba/ rises from 900 to 1240 Hz (Fig. 1). The syllables are identical for the duration of the vowel /a/ (50–170 ms).

The /ba/ and /ga/ stimuli were presented pseudorandomly within the context of 6 other syllables with a probability of occurrence of 12.5% each. The other 6 sounds were also generated using a Klatt-based synthesizer and differed by voice-onset time,  $F_0$  and duration. Syllables were presented in a single block. The recording session lasted  $35 \pm 5$  min. Because we were interested in quantifying the effects of neural discrimination of speech sounds differing only in formant structure, responses to /ga/ and /ba/ are assessed here. See Chandrasekaran et al. (2009) for further descriptions of these other syllables.

#### Procedure

ABRs were differentially recorded at a 20-kHz sampling rate using Ag-AgCl electrodes in a vertical montage (Cz active, FPz ground, and earlobe reference) in Neuroscan Acquire 4.3 (Compumedics, Inc., Charlotte, NC, USA). Contact impedance was  $\leq 5$  k $\Omega$  across electrodes. Stimuli were presented binaurally to adults and monaurally to children at 80 dB SPL with an 83-ms interstimulus interval (Scan 2, Compumedics, Inc.) through insert earphones (ER-3, Etymotic Research, Inc., Elk Grove Village, IL, USA). The speech syllables were presented in alternating polarities; to change a stimulus from one polarity to another, the stimulus waveform was inverted by 180°. Neural responses to each stimulus polarity were subtracted in order to facilitate the analysis of phase-locked neural activity to the frequency range in which the stimuli differed (900–2480 Hz). This technique



**Figure 1.** Stimulus spectrograms (A and B) and cross-phase comparisons (C).

retains the response to the fine structure of the signal while minimizing the envelope response (Aiken and Picton 2008; Akhoun et al. 2008; Skoe and Kraus 2010). During the recording, subjects watched a movie of their choice to facilitate a restful state. Seven hundred artifact-free trials were collected to each stimulus in adults while 850 were collected in both child age groups. Because a subset of adults yielded an extra 100 trials ( $N=19$ ), we present additional analyses in this subgroup on their averaged responses to 800, rather than 700, artifact-free trials in the Results section entitled, “More Distinct Neural Responses to Similar Speech Syllables in Musician Children and Adults.”

ABRs were bandpass filtered offline from 70 to 2000 Hz (12 dB/octave roll-off) to maximize auditory brainstem contributions to the signal and to reduce the influence of low-frequency cortical activity (Akhoun et al. 2008). The lowpass limit was set well above the noise floor of the responses (see Data Analysis section). Responses were baseline-corrected to the prestimulus period (−40 to 0 ms) and trials with amplitudes exceeding  $\pm 35 \mu\text{V}$  were rejected as artifacts. All data processing was conducted using scripts generated in Matlab 2007b (The Mathworks, Natick, MA, USA).

#### Phase Analysis

The cross-phaseogram was constructed according to Skoe et al. (2011). We divided the response into overlapping 20 ms windows starting at 40 ms before the onset of the stimulus, with each window separated from the next by 1 ms. The center time-point of each window ranged from 0 to 170 ms with a total of 170 windows analyzed. These windows were baseline-corrected using Matlab’s detrend function and ramped using a Hanning window. Within each 20 ms window, we applied the Matlab cross-power spectral density function and converted the resulting spectral estimates to phase angles.

Greater positive phase shifts reflect the tonotopic organization of the human auditory system, with sounds composed of higher frequencies eliciting earlier neural responses than sounds comprising lower frequencies (e.g., /ga/ prior to /ba/; see again Skoe et al. 2011). Because of this, positive phase shifts were anticipated over the time-frequency region where the 2 stimuli differed (i.e., over the response to the formant transition and in the frequency range of the second formant; see Fig. 1).

#### Data Analysis

The cross-phaseogram was split into 2 time regions according to the acoustic characteristics of the stimuli: a time region corresponding to the dynamic formant transition (15–45 ms) and another corresponding to the sustained vowel (60–170 ms). We restricted our analyses to the frequency region in which the stimuli differed (900–2450), calculating the upper frequency of the region of interest based on where the neural response fell below the noise floor (defined as prestimulus spectral activity). Whereas the lower limit of the frequency region of interest was 900 Hz for all groups, the upper limit was 1250 Hz for preschoolers, 1400 Hz for adults, and 1500 Hz for school-aged children. The neural response noise floor was determined through visual comparison of the stimulus-evoked response to that of the prestimulus period in the spectral domain via fast Fourier transforms. The noise floor’s lower limit was reached when spectral peaks in individual responses were greater in amplitude than those observed in the prestimulus period for 3 of the 5 highest spectral peaks.

A repeated-measures analysis of variance (RMANOVA) was conducted to assess the phase shifts between responses to /ba/ and /ga/ according to age group and musicianship (as between-subject factors), with time range (corresponding to the formant transition/sustained vowel) as the within-subject factor. Post hoc-independent samples *t*-tests compared the extent of phase shift in musicians and nonmusicians for both time regions for each age group. Because the stimuli were not acoustically distinct during the sustained vowel, this region served as a control; phase shifts and, thus, musician–nonmusician differences (i.e., greater phase shifts in musicians) were only anticipated for the format transition region. We also compared musicians’ and nonmusicians’ cognitive performance using 1-way ANOVAs. Pearson correlations were conducted to explore relationships among the extent of phase shift, cognitive performance, and musical practice histories. Normality for all data was confirmed by the Shapiro–Wilk test. All statistical analyses were performed using SPSS 19 (SPSS, Inc., Chicago, IL, USA) and reflect 2-tailed *P*-values.

## Results

### Summary of Results

Musicians in all 3 age groups demonstrated more temporally distinct responses to /ba/ and /ga/ in the anticipated direction, as indicated by greater positive phase shifts (i.e., responses to /ga/ phase-leading responses to /ba/) in response to the formant transition of the stimuli in the spectral frequency band where the stimuli differ. Musicians also outperformed nonmusicians on auditory attention and working memory. The extent of temporal distinction correlated with auditory but not visual cognitive abilities in school-aged children and adults as well as with years of musical practice in children with the least amount of music training (preschoolers).

### Enhanced Attention and Working Memory in Musician Children and Adults

While IQ did not differ between-groups (see Table 1), musician school-aged children and adults outperformed nonmusicians on auditory but not visual working memory and attention (Table 1; school-aged children: AWM:  $F_{1,25}=6.7$ ,  $P=0.01$ , VWM:  $F_{1,25}=0.3$ ,  $P=0.60$ , AAtt:  $F_{1,25}=4.1$ ,  $P=0.05$ ,

**Table 1**  
Group means (standard deviations) for IQ and cognitive performance in child and adult musicians and nonmusicians

	Preschoolers	School-aged children	Adults
IQ (percentile)			
NonMus	67.3 (26.6)	84.4 (19.3)	63.0 (20.0)
Mus	77.8 (15.3)	89.7 (10.8)	70.6 (18.9)
AWM (score)			
NonMus	—	<b>116.6 (13.4)**</b>	<b>114.1 (11.3)**</b>
Mus	—	<b>128.4 (8.8)</b>	<b>127.2 (13.1)</b>
VWM (digits)			
NonMus	—	4.2 (1.5)	8.5 (1.8)
Mus	—	4.5 (1.5)	8.9 (2.0)
Auditory attention (ms)			
NonMus	—	<b>489.9 (157.7)*</b>	<b>393.1 (82.3)*</b>
Mus	—	<b>391.0 (84.0)</b>	<b>324.2 (46.5)</b>
Visual attention (ms)			
NonMus	—	332.4 (147.9)	316.7 (91.0)
Mus	—	279.6 (110.4)	283.8 (34.3)

Measures by which musicians and nonmusicians differ are indicated in bold (\* $P < 0.05$ , \*\* $P < 0.01$ ).

VAtt:  $F_{1,25} = 1.1$ ,  $P = 0.31$ ; adults AWM:  $F_{1,28} = 8.3$ ,  $P = 0.008$ , VWM:  $F_{1,28} = 0.2$ ,  $P = 0.71$ , AAtt:  $F_{1,16} = 4.9$ ,  $P = 0.04$ , VAtt:  $F_{1,16} = 0.7$ ,  $P = 0.43$ ). Better attention performance was reflected by faster reaction times over the course of the sustained attention task compared with nonmusicians, corroborating previous reports of enhanced attention in musician children and adults (Strait et al. 2010; 2012). Memory and attention tasks were not administered to preschool children.

### More Distinct Neural Responses to Similar Speech Syllables in Musician Children and Adults

A RMANOVA with age group and musicianship (Mus/NonMus) as fixed factors and time range (formant transition/sustained vowel) as within-subject factor revealed a main effect of time range on the extent of phase shift between responses to /ba/ and /ga/ ( $F_{1,70} = 4.9$ ,  $P = 0.03$ ). Furthermore, we observed an interaction between time range and musicianship ( $F_{1,70} = 29.3$ ,  $P < 0.0001$ ). There was no main effect of age or a musicianship  $\times$  age  $\times$  time range interaction, which was expected given that adult and child ABRs were recorded under different parameters (e.g., binaural/monaural, different sweep counts). Post-hoc independent samples  $t$ -tests revealed that musicians have more temporally distinct neural responses to /ba/ and /ga/ corresponding to the formant transitions of the stimuli in the anticipated direction (i.e., greater positive phase shifts) but not to the shared vowel, /a/.

Although adult musicians demonstrated greater positive phase shifts between responses to /ba/ and /ga/ than nonmusicians, nonmusicians' mean phase shift was in the opposite direction, with responses to /ba/ preceding those to /ga/ (Fig. 2, bottom panel). While this reversed pattern may be veridical, it was not observed by Parbery-Clark et al. (2012), in which nonmusicians' phase shifts hover around zero, nor is it reflected in either child age group. Nonmusicians' temporal distinctions between /ba/- and /ga/-evoked responses may not be sufficiently consistent or robust to be captured in comparisons between responses composed of so few sweeps. Because a handful of adults yielded 100 additional artifact-free trials beyond the targeted 700, we reassessed the present data to ask whether the negative phase shifts observed in nonmusicians were a consequence of the small number of individual trials making up the averaged responses

(traditional ABR analyses, such as the measurement of peak timing and magnitudes, are performed on responses comprised of many thousands, rather than hundreds, of trials). We predicted that nonmusicians would trend toward either a lack of phase shifts or shifts in the positive (i.e., anticipated) direction with the addition of more sweeps. Even within this constrained subject population ( $N = 19$ , with Musicians = 10), musicians still demonstrated greater positive phase shifts than nonmusicians to the formant transition region ( $t_{(18)} = -2.1$ ,  $P = 0.05$ ) but not to the vowel ( $t_{(18)} = -1.2$ ,  $P = 0.22$ ). Furthermore, the addition of more trials led musicians' phase shifts to become even greater in the positive direction while the mean nonmusician phase shift moved closer to zero: considering only the subjects included in the 800-trial comparisons, nonmusicians' mean phase shift increased from  $-0.3$  radians with 700 trials to  $-0.18$  radians with 800 trials. In fact, 4 of the 9 nonmusicians now demonstrated phase shifts in the anticipated positive direction while the others gravitated closer toward zero.

### Relationships among Neural Distinction of Similar Speech Syllables, Cognitive Abilities, and Extent of Musical Practice

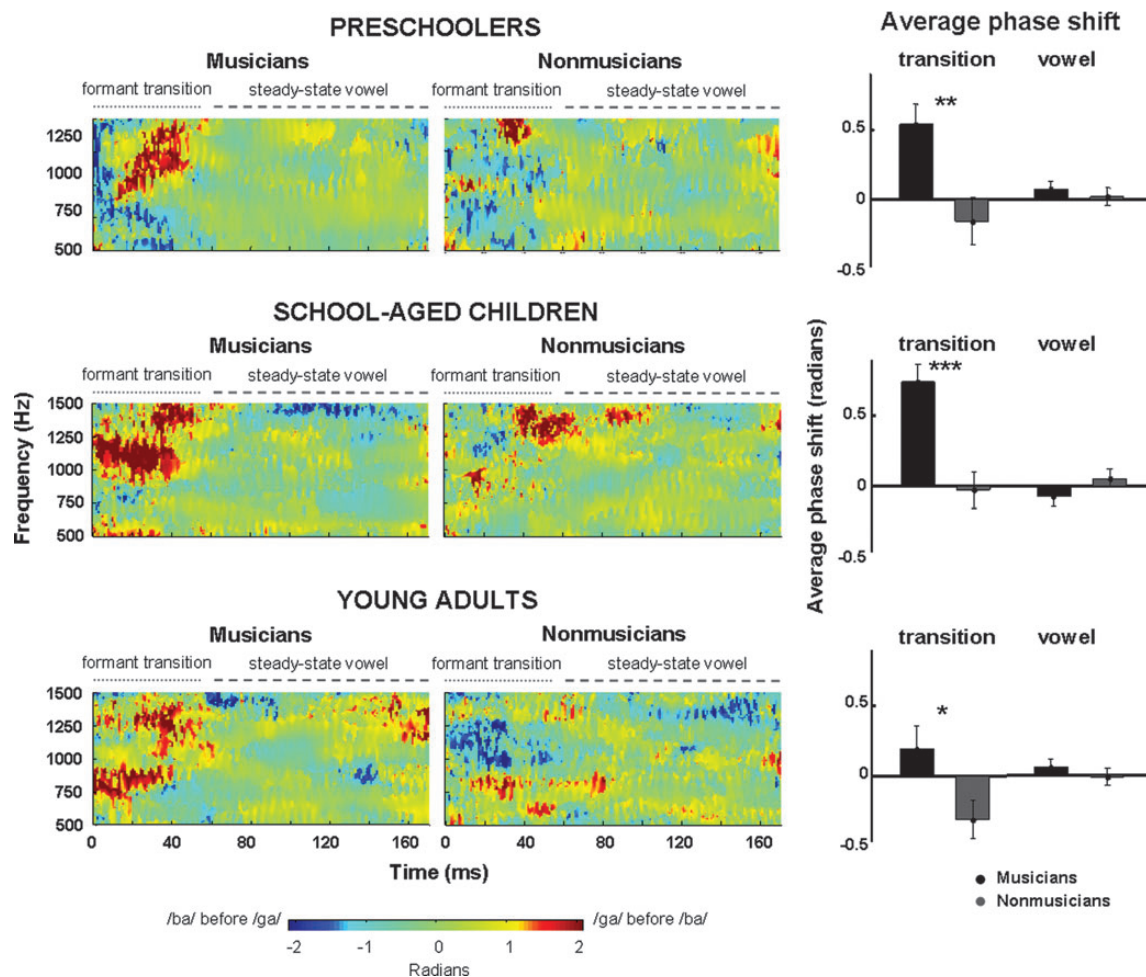
Auditory but not visual cognitive performance correlated with the extent of temporal distinction between neural responses to /ba/ and /ga/ in school-aged children and adults (Fig. 3; Table 2). Although neither years of musical practice nor age of training onset related to the extent of temporal distinction in school-aged children and adults (all  $P \geq 0.4$ ), the total years of music training correlated with phase shifts between responses to the formant transitions of /ba/ and /ga/ in preschoolers (Fig. 4; across all preschoolers: formant transition:  $r = 0.63$ ,  $P = 0.002$ , vowel:  $r = 0.10$ ,  $P = 0.67$ ; limited to preschoolers with some degree of music training (plotted): formant transition:  $r = 0.54$ ,  $P = 0.04$ , vowel:  $r = -0.01$ ,  $P = 0.97$ ). This relationship could not be accounted for by age (age partialled; formant transition:  $r = 0.57$ ,  $P < 0.01$ ; vowel:  $r = 0.04$ ,  $P = 0.87$ ); in fact, age alone did not correlate with phase shifts between responses to either time region (preschoolers; formant transition:  $r = 0.25$ ,  $P = 0.29$ ; vowel:  $r = 0.25$ ,  $P = 0.28$ ).

### Discussion

This study reveals that musicians' more distinct neural encoding of stop consonants arises early in life (i.e., as young as age 3) and is observed in children with as little as a few years of training. Furthermore, musicians' enhanced neural stop consonant distinction correlates with auditory-specific cognitive abilities and, within the youngest children, with extent of musical practice. These relationships may reflect cognitive aspects of musicians' sensory processing enhancements by means of strengthened top-down control over auditory processing.

### Neural Speech-Sound Differentiation in Musicians: Practical Implications

Our results contribute to a growing literature supporting increased neural sensitivity to speech-sound distinctions in musicians. Chobert et al. (2011) recently reported enhanced sensitivity to differences in voice-onset time (i.e., the acoustic

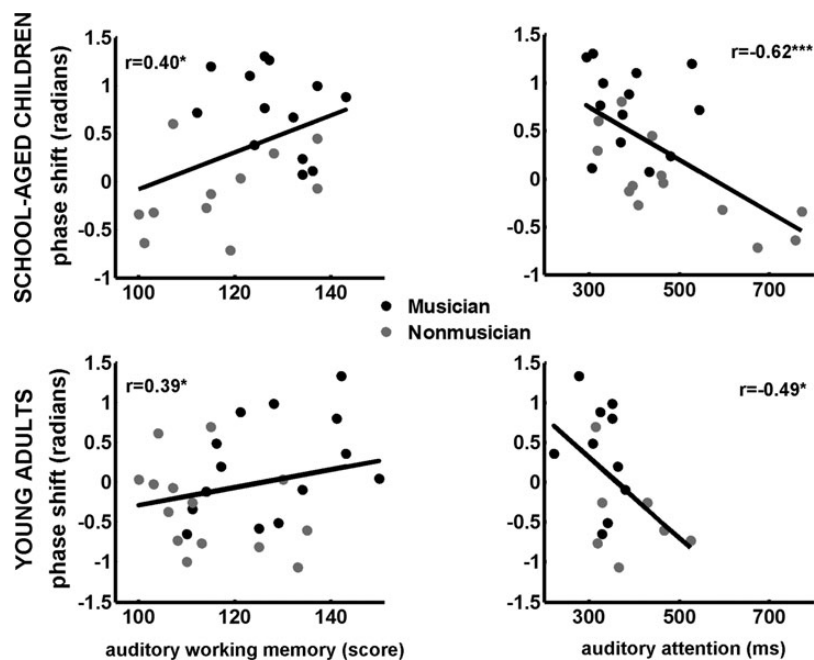


**Figure 2.** Cross-phase comparisons of responses to contrasting speech sounds /ga/ and /ba/ in musicians (left) and nonmusicians (right) for 3 age groups. For each cross-phaseogram, phase differences are plotted as a function of frequency over time. Phase shift in radians is indicated by color, with warm colors indicating phase-lead (/ga/ before /ba/) and cool colors indicating phase-lag (/ba/ before /ga/). Analyses were performed for periods corresponding to the formant transition (15–45 ms) and steady-state vowel (60–170 ms), as notated by dotted lines. Bar graphs indicate average phase shifts for each age group in the frequency range that distinguishes /ba/ from /ga/ up to the noise floor for each age group (adults: 900–1400 Hz, children: 900–1500 Hz, preschoolers: 900–1350 Hz). Musicians have greater phase shifts than nonmusicians in responses to the formant transition. The steady-state vowel is consistent between stimuli and responses show an appropriate lack of phase shift in both musicians and nonmusicians (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ ).

feature distinguishing syllables like da/ta), syllable duration, and vowel identity in 9-year-old musician children and directly linked most of these enhancements to music training using a subsequent longitudinal design (Chobert, Francois, Velay, et al. 2012); we are the first to demonstrate child musician-nonmusician distinctions in the neural processing of speech syllables distinguished by place of articulation (i.e., formant onset frequencies). These functional enhancements may facilitate improved cortical and behavioral indices of speech segmentation, which has been recently reported in musician children using a 2-year longitudinal music training paradigm (Francois et al. 2012), as well as musicians' better language-learning compared with nonmusicians (Shook et al. 2013).

Although we previously established more precise neural encoding of speech syllables in musician adults (Parbery-Clark, Tierney, et al. 2012), the results presented here reveal that this distinction arises in musically trained children as young as 3 years of age—prior to the onset of learning to read and during the years in which the categorical perception

of speech sounds develops (Nittrouer and Miller 1997; Hazan and Barrett 2000). Musical experience prior to 3.5 years of age may be especially influential on speech processing given evidence for the closing of a sensitive period for the development of speech perception at that stage (Zwolan et al. 2004; Harrison et al. 2005; Houston and Miyamoto 2010). School-aged children's enhanced neural speech-sound differentiation may indicate a steepened developmental trajectory for categorical speech perception, providing a biological basis for the impact of musical training on phonological awareness (Dege and Schwarzer 2011), reading acquisition (Moreno et al. 2009), and speech perception (Strait et al. 2012). Although relationships between this neural measure and reading ability and speech-in-noise perception have already been established (Hornickel et al. 2009, 2011; Parbery-Clark, Tierney, et al. 2012), future work should assess such relationships in musician and nonmusician children longitudinally, pre-, and post-musical training onset. Particular attention might be given to preschoolers undergoing music training in order to determine whether the speech perception and



**Figure 3.** Relationships between the subcortical differentiation of speech syllables and auditory cognitive abilities for school-aged children (top) and young adults (bottom). Better performance on tests of auditory working memory (higher scores) and auditory attention (faster reaction times) relates to greater phase differences between responses to the formant transitions of /ga/ and /ba/ (\* $P < 0.05$ , \*\*\* $P < 0.005$ ).

**Table 2**

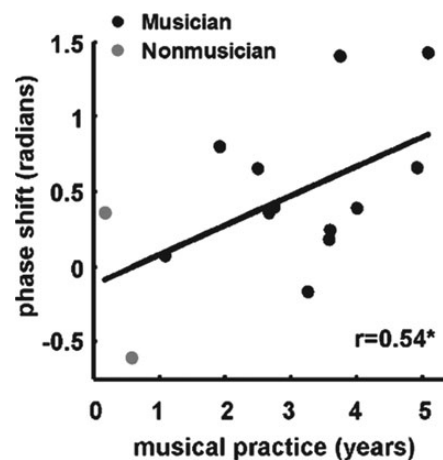
Auditory cognitive abilities correlated with the extent of phase distinctions in neural responses to /ga/ and /ba/ formant transitions in school-aged children and adults

	Phase difference	
	Formant transition	Steady-state vowel
Auditory working memory		
Children	<b>0.40 (&lt;0.05)</b>	0.24 (0.25)
Adults	<b>0.39 (&lt;0.05)</b>	0.18 (0.35)
Visual working memory		
Children	-0.10 (0.64)	-0.15 (0.45)
Adults	-0.08 (0.75)	0.01 (0.97)
Auditory attention		
Children	<b>-0.62 (&lt;0.001)</b>	-0.31 (0.12)
Adults	<b>-0.49 (&lt;0.05)</b>	0.20 (0.45)
Visual attention		
Children	-0.09 (0.66)	0.02 (0.94)
Adults	-0.16 (0.61)	0.08 (0.81)

Values depict Pearson's  $r$  ( $P$ -value), with significant correlations indicated in bold.

literacy-related advantages reported in school-aged child musicians (e.g., phonological awareness) emerge during early childhood.

Although adult musicians demonstrate greater positive phase shifts between responses to /ba/ and /ga/ than nonmusicians, nonmusicians' mean phase shift is of equivalent magnitude in the opposite direction, with responses to /ba/ preceding those to /ga/. It is possible that this outcome does not accurately reflect nonmusicians' temporal distinctions between responses to these speech syllables; nonmusicians' temporal distinctions may not be sufficiently robust or consistent to capture them in comparisons of responses comprising so few individual trials. This contrast may stem from less consistent responses in nonmusicians (Parbery-Clark, Anderson, et al. 2012b), which could necessitate the collection of more trials. Given that greater positive phase shifts reflect the



**Figure 4.** Degree of subcortical differentiation in responses to /ba/ and /ga/ correlates with preschoolers' extent of musical practice (\* $P < 0.05$ ).

tonotopic organization of the human auditory system, with sounds composed of higher frequencies eliciting earlier neural responses than sounds comprising lower frequencies (e.g., /ga/ prior to /ba/), we do not know what advantage the opposite relationship would yield. Further work might reassess this neural metric in young adults as well as in school-aged and preschool children to determine whether increasing the number of stimulus repetitions even beyond the additional analysis presented here in adults yields the measurement of positive phase shifts in nonmusicians, with responses to /ga/ phase-leading responses to /ba/. In this case, we would still predict musicians to demonstrate more temporally distinct neural responses than nonmusicians, with responses to /ga/ preceding responses to /ba/ to a greater extent.

### ***Innate Characteristics or Training-Related Outcomes?***

A growing body of work points to the connectedness of music and language abilities independent of musical training, including reading ability (Anvari et al. 2002; Overy et al. 2003; Slevc and Miyake 2006; Strait et al. 2011), speech-sound differentiation (Anvari et al. 2002; Milovanov et al. 2009), and learning a foreign language (Slevc and Miyake 2006; Milovanov et al. 2008; Milovanov et al. 2010). This raises the question: does musicians' enhanced speech processing reflect inherent differences between individuals with and without musical propensities? Although we interpret differences between musicians and nonmusicians to reflect training-related outcomes, our cross-sectional paradigm cannot directly disentangle training-related from innate characteristics of musicians' auditory brainstem and cognitive function.

It is clear that a confluence of innate and training-related factors shape ABR characteristics. Speech-evoked neural responses have the capacity to reflect genetic components (Hornickel et al. 2013), which may lead one (or, in the case of our youngest participants, one's parent) to pursue musical training. Innate differences, however, cannot account for relationships between the precision of speech-evoked brainstem responses and years of music practice, seen here in our youngest participants and reported elsewhere in school-aged children and adults (Musacchia et al. 2007; Wong et al. 2007; Parbery-Clark et al. 2009; Strait et al. 2009; Skoe and Kraus 2012). In fact, a recent study not only indicates that the neural benefits of music training relate to years of music practice but further demonstrates their perseverance even after many years of discontinued practice (Skoe and Kraus 2012). Relationships such as these suggest that musicians' subcortical advantages reflect training-related contributors, to some extent.

Our previous report in adults did not reveal a relationship between extent of music training and neural response differences to these same speech syllables (Parbery-Clark, Tierney, et al. 2012). We consider the possibility that the lack of this relationship could reflect the highly trained nature of the population. If the benefits accrued from music practice are most significant during the first few years of training, a group of individuals with minimal-to-moderate amounts of musical experience might be necessary to yield a correlation between years of practice and neural speech-sound distinction. The present study demonstrates such a relationship in our least-trained age group. This correlation sheds light on the malleability of the neural mechanism itself: as little as a few years of training may engender significant enhancements in the subcortical differentiation of closely related speech syllables. Although continued practice beyond the first few years may or may not be required to sustain musicians' speech-sound processing advantage according to this mechanism, it may not engender further enhancement.

### ***Subcortical Auditory Function Reflects Cognitive Processes***

Rapidly accruing evidence indicates that the ABR reflects cognitive contributors to auditory processing: faster and more robust ABRs relate to more proficient AWM and attention (Kraus et al. 2012; Krizman et al. 2012; Strait et al. 2012). Here, we demonstrate these relationships across 2 developmental populations: in both school-aged children and adults,

the extent of neural distinction observed to similar speech syllables correlates with auditory but not visual working memory and attention. Although we recognize that our methodology cannot define the mechanisms by which subcortical and cognitive aspects of auditory function converse, the consistency of cognitive-brainstem interactions over development provides a window into the working circuitry of the auditory system, in which sensory function cannot be divorced from higher level influences.

These results can be considered in the context of the modulatory characteristics of the mammalian descending auditory system. A vast literature substantiates direct contributions of cortical centers of memory and attention on basic auditory function (Weinberger 2004; Fritz, Elhilali, David, et al. 2007; Fritz, Elhilali, Shamma 2007). Top-down effects are facilitated by functional connections between extra-sensory cortices associated with executive function (e.g., prefrontal, anterior cingulate cortex cortices) and primary auditory cortex (Pandya et al. 1981; Morris et al. 1998; Yan and Zhang 2005; Crottaz-Herbette and Menon 2006; Fritz et al. 2010). Descending cortico-thalamic, -collicular, and -cochlear nuclear projections mediate effects at lower level nuclei (Yan et al. 2005; Luo et al. 2008; Bauerle et al. 2011), adding to modulatory influences by limbic centers involved in memory and attention (Marsh et al. 2002; Macedo et al. 2005) and the cholinergic system (Zhang et al. 2005; Ji and Suga 2009). In fact, descending connections are crucial in order for learning to shape subcortical sensory function (Bajo et al. 2010; Bauerle et al. 2011). We propose that the correlations we observe among auditory cognitive abilities and subcortical response characteristics reflect strengthened corticofugal systems in individuals with more precise speech-evoked ABRs, regardless of age or life experience.

In addition to providing a window into cognitive-brainstem interactions, our results may reflect functional and anatomical impacts of musical training on the development of the descending auditory system during childhood. Although myelination of the human auditory nerve and brainstem is thought to be complete within the first 2 years of life (Moore et al. 1995; Moore and Linthicum 2007), structural and functional organization of auditory cortex continues well into adolescence, with deeper layers reaching adult-like states prior to more superficial layers (Gleeson and Walsh 2000; Moore and Guan 2001; Moore and Linthicum 2007). The subsequent maturation of layers I–III, which receive input from other cortical sites, underscores the later emergence of the cortico-cortical connectivity necessary for top-down influences to shape sensory function. This ongoing cortical maturation may account for developmental changes in speech-evoked auditory brainstem function that extend well into childhood (Johnson et al. 2008) by means of the descending auditory system. Whereas the development of some aspects of bottom-up auditory processing appear to be predetermined, occurring in the absence of auditory stimulation (Taniguchi 1981), the later emergence of the descending system is more dependent on interactions with the auditory environment, especially during the sensitive developmental period comprising the first 4–9 years of life (for review, see Kral and Eggermont (2007)). Although acoustically enriched environments without behavioral relevance lead to habituated auditory-evoked responses (Norena et al. 2006), acoustic enrichment paired with behavioral relevance increases evoked response amplitudes and neuronal firing

rates (Engineer et al. 2004). Musical training throughout this time period may strengthen the development of top-down auditory connectivity, decreasing cell death and increasing synaptic density between auditory processing sites, given music's inherently rewarding characteristics.

## Conclusions

Here, we present evidence for greater subcortical differentiation of stop consonants in musician children as young as age 3, during a sensitive period for auditory development. Given relationships between subcortical speech-sound distinctions and critical language and reading skills, music training may offer an efficient means of improving auditory processing in young children. We propose that music training provides a rewarding, acoustically enriched auditory environment during this sensitive developmental period, with the capacity to confer lasting impacts on the functional and structural organization of the human auditory system that set the stage for later language skills.

## Funding

This work was supported by the National Institutes of Health (grant F31DC011457-01 to D.S.), the National Science Foundation (grant 0921275 to N.K.), and the Knowles Hearing Center.

## Notes

The authors thank Karen Chan, Victor Abecassis, and Emily Hittner for their assistance with data collection and Trent Nicol, Travis White-Schwoch, and Jennifer Krizman for their comments on the manuscript. *Conflict of Interest:* None declared.

## References

Aiken SJ, Picton TW. 2008. Envelope and spectral frequency-following responses to vowel sounds. *Hear Res.* 245:35–47.

Akhoun I, Gallego S, Moulin A, Menard M, Veuillet E, Berger-Vachon C, Collet L, Thai-Van H. 2008. The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clin Neurophysiol.* 119:922–933.

Anvari SH, Trainor LJ, Woodside J, Levy BA. 2002. Relations among musical skills, phonological processing, and early reading ability in preschool children. *J Exp Child Psychol.* 83:111–130.

Bajo VM, Nodal FR, Moore DR, King AJ. 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci.* 13:253–260.

Banai K, Hornickel JM, Skoe E, Nicol T, Zecker S, Kraus N. 2009. Reading and subcortical auditory function. *Cereb Cortex.* 19:2699–2707.

Barry JG, Ferguson MA, Moore DR. 2010. Making sense of listening: the IMAP test battery. *J Vis Exp.* 44.

Bauerle P, von der Behrens W, Kossel M, Gaese BH. 2011. Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *J Neurosci.* 31:9708–9722.

Bidelman GM, Gandour JT, Krishnan A. 2009. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci.* 23:425–434.

Bidelman GM, Krishnan A. 2010. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* 1355:112–125.

Billiet CR, Bellis TJ. 2011. The relationship between brainstem temporal processing and performance on tests of central auditory function in children with reading disorders. *J Speech Lang Hear Res.* 54:228–242.

Bradley L, Bryant PE. 1983. Categorising sounds and learning to read: a causal connection. *Nature.* 310:419–421.

Brown L, Sherbenou RJ, Johnsen SK. 1997. Test of Nonverbal Intelligence (TONI-3). Austin (TX): Pro-Ed, Inc.

Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N. 2009. Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron.* 64:311–319.

Chobert J, Francois C, Habib M, Besson M. 2012. Deficit in the preattentive processing of syllabic duration and VOT in children with dyslexia. *Neuropsychologia.* 50:2044–2055.

Chobert J, Francois C, Velay JL, Besson M. 2012. Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cereb Cortex.* doi: 10.1093/cercor/bhs377.

Chobert J, Marie C, Francois C, Schon D, Besson M. 2011. Enhanced passive and active processing of syllables in musician children. *J Cogn Neurosci.* 23:3874–3887.

Crottaz-Herbette S, Menon V. 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *J Cogn Neurosci.* 18:766–780.

Davis HP, Keller F. 2002. Colorado assessment tests (CATS), version 1.2. Colorado Springs (CO): Colorado Assessment Tests.

Dege F, Schwarzer G. 2011. The effect of a music program on phonological awareness in preschoolers. *Front Psychol.* 2:124.

Dunn LM, Dunn LM. 1997. Peabody picture vocabulary test. 3rd ed. San Antonio (TX): Pearson Education, Inc.

Engineer ND, Percaccio CR, Pandya PK, Moucha R, Rathbun DL, Kilgard MP. 2004. Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons. *J Neurophysiol.* 92:73–82.

Francois C, Chobert J, Besson M, Schon D. 2012. Music training for the development of speech segmentation. *Cereb Cortex.* (Epub ahead of print). doi: 10.1093/cercor/bhs180.

Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA. 2010. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat Neurosci.* 13:1011–1019.

Fritz JB, Elhilali M, David SV, Shamma SA. 2007. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear Res.* 229:186–203.

Fritz JB, Elhilali M, Shamma SA. 2007. Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J Neurophysiol.* 98:2337–2346.

Gleeson JG, Walsh CA. 2000. Neuronal migration disorders: from genetic diseases to developmental mechanisms. *Trends Neurosci.* 23:352–359.

Harrison RV, Gordon KA, Mount RJ. 2005. Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. *Dev Psychobiol.* 46:252–261.

Hazan V, Barrett S. 2000. The development of phonemic categorization in children aged 6–12. *J Phonetics.* 28:377–396.

Hornickel J, Chandrasekaran B, Zecker S, Kraus N. 2011. Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. *Behav Brain Res.* 216:597–605.

Hornickel J, Lin D, Kraus N. 2013. Speech-evoked auditory brainstem responses reflect familial and cognitive influences. *Dev Sci.* 16:101–110.

Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. 2009. Subcortical differentiation of voiced stop consonants: relationships to reading and speech in noise perception. *Proc Natl Acad Sci USA.* 106:13022–13027.

Houston DM, Miyamoto RT. 2010. Effects of early auditory experience on word learning and speech perception in deaf children with cochlear implants: implications for sensitive periods of language development. *Otol Neurotol.* 31:1248–1253.



- Ji W, Suga N. 2009. Tone-specific and nonspecific plasticity of inferior colliculus elicited by pseudo-conditioning: role of acetylcholine and auditory and somatosensory cortices. *J Neurophysiol.* 102: 941–952.
- Johnson KL, Nicol T, Zecker SG, Kraus N. 2008. Developmental plasticity in the human auditory brainstem. *J Neurosci.* 28:4000–4007.
- Klatt D. 1980. Software for a cascade/parallel formant synthesizer. *J Acoust Soc Amer.* 67:13–33.
- Kral A, Eggermont JJ. 2007. What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res Rev.* 56:259–269.
- Kraus N, Chandrasekaran B. 2010. Music training for the development of auditory skills. *Nat Rev Neurosci.* 11:599–605.
- Kraus N, McGee T, Carrell TD, Sharma A. 1995. Neurophysiologic bases of speech discrimination. *Ear Hear.* 16:19–37.
- Kraus N, McGee TJ, Carrell TD, Zecker SG, Nicol TG, Koch DB. 1996. Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science.* 273:971–973.
- Kraus N, Strait DL, Parbery-Clark A. 2012. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann N Y Acad Sci.* 1252:100–107.
- Krizman J, Marian V, Shook A, Skoe E, Kraus N. 2012. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci USA.* 109: 7877–7881.
- Luo F, Wang Q, Kashani A, Yan J. 2008. Corticofugal modulation of initial sound processing in the brain. *J Neurosci.* 28:11615–11621.
- Macedo CE, Cuadra G, Molina V, Brandao ML. 2005. Aversive stimulation of the inferior colliculus changes dopamine and serotonin extracellular levels in the frontal cortex: modulation by the basolateral nucleus of amygdala. *Synapse.* 55:58–66.
- Marsh RA, Fuzessery ZM, Grose CD, Wenstrup JJ. 2002. Projection to the inferior colliculus from the basal nucleus of the amygdala. *J Neurosci.* 22:10449–10460.
- Milovanov R, Huotilainen M, Esquef PA, Alku P, Valimaki V, Tervaniemi M. 2009. The role of musical aptitude and language skills in preattentive duration processing in school-aged children. *Neurosci Lett.* 460:161–165.
- Milovanov R, Huotilainen M, Valimaki V, Esquef PA, Tervaniemi M. 2008. Musical aptitude and second language pronunciation skills in school-aged children: neural and behavioral evidence. *Brain Res.* 1194:81–89.
- Milovanov R, Pietila P, Tervaniemi M, Esquef PA. 2010. Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education. *Learn Individ Diff.* 20:56–60.
- Moore JK, Guan YL. 2001. Cytoarchitectural and axonal maturation in human auditory cortex. *J Assoc Res Otolaryngol.* 2:297–311.
- Moore JK, Linthicum FH Jr. 2007. The human auditory system: a timeline of development. *Int J Audiol.* 46:460–478.
- Moore JK, Perazzo LM, Braun A. 1995. Time course of axonal myelination in the human brainstem auditory pathway. *Hear Res.* 87:21–31.
- Moreno S, Marques C, Santos A, Santos M, Castro SL, Besson M. 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb Cortex.* 19:712–723.
- Morris JS, Friston KJ, Dolan RJ. 1998. Experience-dependent modulation of tonotopic neural responses in human auditory cortex. *Proc Biol Sci.* 265:649–657.
- 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 USA.* 104:15894–15898.
- Nittrouer S, Miller ME. 1997. Predicting developmental shifts in perceptual weighting schemes. *J Acoust Soc Am.* 101:2253–2266.
- Norena AJ, Gourevitch B, Aizawa N, Eggermont JJ. 2006. Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. *Nat Neurosci.* 9:932–939.
- Overy K, Nicolson RI, Fawcett AJ, Clarke EF. 2003. Dyslexia and music: measuring musical timing skills. *Dyslexia.* 9:18–36.
- Pandya DN, Van Hoesen GW, Mesulam MM. 1981. Efferent connections of the cingulate gyrus in the rhesus monkey. *Exp Brain Res.* 42:319–330.
- Parbery-Clark A, Anderson S, Hittner E, Kraus N. 2012a. Musical experience offsets age-related delays in neural timing. *Neurobiol Aging.* 33:1483.e1. –1483.e4.
- Parbery-Clark A, Anderson S, Hittner E, Kraus N. 2012b. Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front Aging Neurosci.* 4:1–12.
- Parbery-Clark A, Skoe E, Kraus N. 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci.* 29:14100–14107.
- Parbery-Clark A, Strait DL, Kraus N. 2011. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia.* 49:3338–3345.
- Parbery-Clark A, Tierney A, Strait DL, Kraus N. 2012. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience.* 219:111–119.
- Patel AD. 2011. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front Psychol.* 2:142.
- Rocha-Muniz CN, Befi-Lopes DM, Schochat E. 2012. Investigation of auditory processing disorder and language impairment using the speech-evoked auditory brainstem response. *Hear Res.* 294: 143–152.
- Shook A, Marian V, Bartolotti J, Schroeder S. 2013. Musical experience influences novel language learning. *Am J Psychol.* 126: 95–104.
- Siok WT, Fletcher P. 2001. The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Dev Psychol.* 37:886–899.
- Skoe E, Kraus N. 2010. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31(3):302–324.
- Skoe E, Kraus N. 2012. A little goes a long way: how the adult brain is shaped by musical training in childhood. *J Neurosci.* 32: 11507–11510.
- Skoe E, Nicol T, Kraus N. 2011. Cross-phaseogram: objective neural index of speech sound differentiation. *J Neurosci Methods.* 196:308–317.
- Slevc LR, Miyake A. 2006. Individual differences in second-language proficiency: does musical ability matter? *Psychol Sci.* 17:675–681.
- Stevens C, Lauinger B, Neville H. 2009. Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. *Dev Sci.* 12:634–646.
- Strait DL, Hornickel J, Kraus N. 2011. Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav Brain Funct.* 7:44.
- Strait DL, Kraus N. 2011. Playing music for a smarter ear: cognitive, perceptual and neurobiological evidence. *Music Percept.* 29:133–147.
- 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.
- Strait DL, Kraus N, Skoe E, Ashley R. 2009. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *Eur J Neurosci.* 29:661–668.
- 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.
- Taniguchi I. 1981. Plastic changes in the inferior colliculus following cochlear destruction. In: Syka J, Aitkin L, editors. *Neuronal mechanisms of hearing.* New York (NY): Plenum Press. p. 377–380.
- Wechsler D. 1999. Wechsler abbreviated scale of intelligence (WASI). San Antonio (TX): Harcourt Assessment.
- Weinberger NM. 2004. Specific long-term memory traces in primary auditory cortex. *Nat Rev Neurosci.* 5:279–290.
- 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.

- Woodcock RW, McGre KS, Mather N. 2001. Woodcock-Johnson psycho-educational battery. 3rd edn. Itasca (IL): Riverside.
- Yan J, Zhang Y. 2005. Sound-guided shaping of the receptive field in the mouse auditory cortex by basal forebrain activation. *Eur J Neurosci.* 21:563–576.
- Yan J, Zhang Y, Ehret G. 2005. Corticofugal shaping of frequency tuning curves in the central nucleus of the inferior colliculus of mice. *J Neurophysiol.* 93:71–83.
- Zhang Y, Hakes JJ, Bonfield SP, Yan J. 2005. Corticofugal feedback for auditory midbrain plasticity elicited by tones and electrical stimulation of basal forebrain in mice. *Eur J Neurosci.* 22:871–879.
- Zwolan TA, Ashbaugh CM, Alarfaj A, Kileny PR, Arts HA, El-Kashlan HK, Telian SA. 2004. Pediatric cochlear implant patient performance as a function of age at implantation. *Otol Neurotol.* 25:112–120.