



## Musical training during early childhood enhances the neural encoding of speech in noise

Dana L. Strait<sup>a,b</sup>, Alexandra Parbery-Clark<sup>a,c</sup>, Emily Hittner<sup>a,d</sup>, Nina Kraus<sup>a,b,c,d,e,f,\*</sup>

<sup>a</sup> Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL, USA

<sup>b</sup> Institute for Neuroscience, Northwestern University, Evanston, IL, USA

<sup>c</sup> Communication Sciences, Northwestern University, Evanston, IL, USA

<sup>d</sup> Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL, USA

<sup>e</sup> Neurobiology and Physiology, Northwestern University, Evanston, IL, USA

<sup>f</sup> Otolaryngology, Northwestern University, Evanston, IL, USA

### ARTICLE INFO

#### Article history:

Accepted 9 September 2012

Available online 25 October 2012

#### Keywords:

Auditory  
Brainstem  
ABR  
Speech in noise  
Attention  
Memory  
Musicians  
Children  
Development

### ABSTRACT

For children, learning often occurs in the presence of background noise. As such, there is growing desire to improve a child's access to a target signal in noise. Given adult musicians' perceptual and neural speech-in-noise enhancements, we asked whether similar effects are present in musically-trained children. We assessed the perception and subcortical processing of speech in noise and related cognitive abilities in musician and nonmusician children that were matched for a variety of overarching factors. Outcomes reveal that musicians' advantages for processing speech in noise are present during pivotal developmental years. Supported by correlations between auditory working memory and attention and auditory brainstem response properties, we propose that musicians' perceptual and neural enhancements are driven in a top-down manner by strengthened cognitive abilities with training. Our results may be considered by professionals involved in the remediation of language-based learning deficits, which are often characterized by poor speech perception in noise.

© 2012 Elsevier Inc. All rights reserved.

### 1. Introduction

Although hearing speech in noisy environments is difficult for everyone, children are particularly vulnerable to the deleterious effects of background noise (Elliott, 1979; Fallon, Trehub, & Schneider, 2000; Hall, Grose, Buss, & Dev, 2002). Perceptual difficulty in noise has been associated with delayed neural timing and decreased encoding of the spectral cues of speech, relative to neural responses to speech in quiet listening conditions (Anderson, Chandrasekaran, Skoe, & Kraus, 2010; Song, Skoe, & Banai, 2010). Noise is especially problematic for children with auditory-based learning impairments, including specific language impairment (Ziegler, Pech-Georgel, George, Alario, & Lorenzi, 2005), developmental dyslexia (Sperling, Lu, Manis, & Seidenberg, 2005; Ziegler, Pech-Georgel, George, & Lorenzi, 2009) and auditory processing disorders (Muchnik et al., 2004). Given that learning often occurs in noisy environments, accurate speech perception in noise is a critical component of early childhood education. We do not know,

however, how intensive auditory training during childhood, such as *musical* training, impacts brain systems that underlie hearing in noise, despite evidence to that effect in adults (Parbery-Clark, Skoe, & Kraus, 2009).

Musical training pervasively impacts aspects of brain structure (Bangert & Schlaug, 2006; Hutchinson, Lee, Gaab, & Schlaug, 2003; Schlaug, 2001), function (Besson, Schon, Moreno, Santos, & Magne, 2007; Chobert, Marie, Francois, Schon, & Besson, 2011; Gaab & Schlaug, 2003; Magne, Schon, & Besson, 2006; Seppanen, Brattico, & Tervaniemi, 2007; Seppanen, Pesonen, & Tervaniemi, 2012; Tervaniemi et al., 2009) and development (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Hyde et al., 2009; Moreno et al., 2009, 2011; Schlaug, Norton, Overy, & Winner, 2005; Shahin, Roberts, & Trainor, 2004), with particular influence on the auditory system (Kraus & Chandrasekaran, 2010; Strait & Kraus, 2011b). Enhancements in musicians have been observed in a structure as evolutionarily ancient as the auditory brainstem, with adult musicians demonstrating more robust subcortical processing of both music and speech relative to nonmusicians (Musacchia, Sams, Skoe, & Kraus, 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007). Subcortical speech processing enhancements are most evident in young adult musicians in challenging perceptual environments, with musicians demonstrating more robust subcortical representation of speech

\* Corresponding author at: Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA.

E-mail address: [nkraus@northwestern.edu](mailto:nkraus@northwestern.edu) (N. Kraus).

URL: <http://www.brainvolts.northwestern.edu> (N. Kraus).

harmonics and less response degradation in adverse listening conditions compared to nonmusicians (Bidelman & Krishnan, 2010; Parbery-Clark et al., 2009). It has been proposed that musicians' neural enhancements for auditory processing reflect their enhanced auditory cognitive abilities. Strengthened auditory working memory and attention (Chan, Ho, & Cheung, 1998; Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011; Strait & Kraus, 2011a; Strait, Kraus, Parbery-Clark, & Ashley, 2010), for example, may sharpen musicians' neural encoding of sound in a top-down fashion (Kraus & Chandrasekaran, 2010; Strait & Kraus, 2011b; Kraus, Parbery-Clark, & Strait, 2012). Whereas studies have addressed how music lessons during childhood impact auditory cognitive development, nothing is known about how musical training shapes the neural processing of speech in noise and related perceptual abilities in children.

Here, we aimed to define associations between musical training during early childhood and the subcortical encoding of speech in noise, speech-in-noise perception and memory and attention, which are cognitive contributors to speech-in-noise processing (Conway, Cowan, & Bunting, 2001; Parbery-Clark et al., 2009, 2011; Strait & Kraus, 2011a). We asked if musically-trained children have enhanced resiliency in the subcortical encoding of speech by assessing the auditory brainstem encoding of speech in quiet and noisy backgrounds. Perceptual and cognitive abilities were assessed using standardized measures of speech-in-noise perception, IQ, attention and working memory. We hypothesized that musically-trained children demonstrate more robust subcortical representation of the temporal and spectral components of speech, which lend to less neural response degradation in noise. We further hypothesized that these enhancements relate with perceptual and auditory-specific cognitive abilities. We anticipated that musicians would demonstrate less timing delays with the addition of background noise in responses to the formant transition, the most spectrotemporally-dynamic, informationally-salient and perceptually-vulnerable portion of the speech stimulus (Tallal & Stark, 1981; Van Tasell, Hagen, Koblas, & Penner, 1982).

## 2. Materials and methods

### 2.1. Participants

All experimental procedures were approved by the Northwestern University Institutional Review Board. Thirty-one normal hearing children ( $\leq 20$  dB pure tone thresholds at octave frequencies from 125 to 8000 Hz) between the ages of 7–13 participated in this study ( $M = 10.2$ ,  $SD = 1.8$  years), for which legal guardians and participants provided informed consent and assent, respectively. Inclusionary criteria also included normal wave V click-evoked ABR latencies. Parents completed an extensive questionnaire addressing the participant's family history, musical experience, extracurricular involvement and educational history. Musicians (Mus,  $N = 15$ ) were self-categorized, were currently undergoing private instrumental training, began musical training by age 5 (mean = 2.0,  $SD = 1.4$ ) and had consistently practiced for at least 4 years (consistency defined as practicing  $\geq 20$  min at least 5 days weekly) (mean = 7.9,  $SD = 2.2$ ). Nonmusicians (NonMus,  $N = 16$ ) were self-categorized and had  $< 5$  years of accumulated musical experience throughout their lifespans (mean = 1.2,  $SD = 1.8$ ). Three of the 16 NonMus participants had some degree of previous musical experience, including preschool music programs (e.g., Kindermusik and Wiggleworms classes). Of these participants, one engaged in a weekly preschool music program (starting at age 1, for 4.5 years of total involvement). The other two had temporarily engaged in private music lessons (one for 1 year, the other for a

half-year). The 13 remaining NonMus participants had no musical experience. Mus and NonMus groups did not differ according to age ( $F_{(1,29)} = 0.03$ ,  $p = 0.86$ ; Mus mean = 10.3 years,  $SD = 1.6$ ; NonMus mean = 10.1 years,  $SD = 1.9$ ), sex ( $\chi^2 = 0.23$ ,  $p = 0.30$ ), IQ (as measured by the 2-subtest Wechsler Abbreviated Scale of Intelligence, comprised of vocabulary and matrix reasoning subtests;  $F_{(1,29)} = 0.23$ ,  $p = 0.63$ ; Harcourt Assessment, San Antonio, TX), by extent of extracurricular activity involvement (measured in average hours per year since birth;  $F_{(1,29)} = 0.997$ ,  $p = 0.33$ ), or by socioeconomic status as inferred from maternal education ( $F_{(1,29)} = 0.32$ ,  $p = 0.58$ ) (see Stevens, Lauinger, & Neville, 2009, for discussion regarding the predictive value of maternal education for inferring a child's socioeconomic status). Nonmusicians' extracurricular activities included chess club, theater-related activities and athletics, among others.

### 2.2. Speech-in-noise (SIN), attention and working memory performance

#### 2.2.1. Speech in noise

SIN perception was measured in a soundproof booth using two standardized measures that varied with respect to the amount of contextual cues conveyed in the target signal: the Words in Noise Test (WIN) and the Hearing in Noise Test (HINT).

WIN is a non adaptive test of SIN perception in four-talker babble noise (Wilson, 1993). Single words are presented through a loudspeaker placed one meter in front of participants. Participants are asked to repeat the words, one at a time. Thirty-five words are presented at 70 dB SPL with a starting signal-to-noise ratio (SNR) of 24 dB, decreasing in 4 dB steps until 0 dB with five words presented at each SNR. Each subject's threshold was based on the number of correctly repeated words, with a lower score indicating better performance.

HINT is an adaptive test of SIN processing in which participants repeat sentences presented in speech-shaped noise from a loudspeaker located one meter directly ahead (Nilsson, Soli, & Sullivan, 1994). The listener has access to acoustic, syntactic and semantic cues that increase the probability of selecting the correct target word from like-sounding competitors. Because of this, HINT performance does not solely rely on peripheral hearing function but also depends on cognitive skills, such as auditory working memory (Parbery-Clark et al., 2009, 2011) and attention (Strait & Kraus, 2011a). Participants completed three conditions across which the location of the noise varied. For the HINT<sub>front</sub> condition, the competing noise was presented from the same loudspeaker as the target sentences (0° azimuth), while for the two subsequent conditions the target sentences continued to be presented via the loudspeaker in front of the participant but the noise was presented at  $-90^\circ$  or  $+90^\circ$  azimuth (HINT<sub>left/right</sub>). For all conditions, the noise presentation level was fixed at 65 dB SPL and the program adjusted the perceptual difficulty by increasing or decreasing the intensity level of the target sentences until 50% of sentences are correctly repeated. This threshold, in dB, was normed according to age (in years), generating percentile rankings. HINT<sub>right</sub> and HINT<sub>left</sub> percentile ranks were averaged to generate a HINT<sub>left/right</sub> composite score.

Because of its reliance on contextual cues and employment of longer stimuli, HINT involves cognitive processes to a greater extent than WIN. WIN, in contrast, is more reflective of peripheral aspects of hearing function in noise (e.g., hearing thresholds) (Parbery-Clark et al., 2011; Wilson, McArdle, & Smith, 2007). Still, perceiving and repeating words in noise is not a purely perceptual task; although less dependent on cognitive processes than HINT, it reflects aspects of language knowledge such as vocabulary and phonotactics (Flege, Meador, & MacKay, 2000).

### 2.2.2. Attention

Auditory and visual attention were assessed using the *Integrated Visual and Auditory Plus Continuous Performance Test* (Sandford & Turner, 1994). The test was administered in a soundproof booth on a laptop computer placed 60 cm from the participant and was divided into four sections: warm-up, practice, test and cool-down. Participants were instructed to use an external mouse to “play the computer game”; further instructions came from the test via child-sized Sennheiser HD 25-1 headphones and corresponding visual cues. During the warm-up, participants were instructed to click the mouse when they saw or heard a “1”; the test proceeded with a 20-trial warm-up during which only the number “1” was spoken or presented visually, 10 times each. Next, participants completed a practice session during which they were reminded of the same instructions but were also asked *not* to click the mouse when they saw or heard a “2”; further practice trials were presented (10 auditory and 10 visual targets). During the main test portion of the test, choice reaction time was recorded for participants’ responses to the target (“1”) and foil (“2”) stimuli on five sets of 100 trials for a total of 500 trials. Each set consisted of two blocks of 50 trials each, with each trial lasting 1.5 s. The visual targets were presented for 167 ms and were 4 cm high, while the auditory stimuli lasted 500 ms and were spoken by a female.

The first block of each set of the main test collects a measure of impulsivity by creating a ratio of target to foil of 5.25:1.0, resulting in 84% of trials (or 42 out of 50, per block) presenting targets intermixed with eight foils. The second block collects a measure of inattention by reversing the order and presenting many foils and few targets (165 targets over all five sets). Stimuli are presented in a pseudo-random order of visual and auditory stimuli. The test is followed by a “cool down,” which mimics the initial practice period; performance on the warm up is compared to practice performance. The duration of the main portion of the test is 13 min, although the entire assessment including the introduction, practice, test and cool down lasts 20 min. The assessment generates a primary diagnostic scale called the “full scale attention quotient” score, which can be divided into auditory and visual components.

### 2.2.3. Working memory

Auditory and visual working memory (AWM and VWM) were assessed using the *Woodcock Johnson III Test of Cognitive Abilities* Auditory Working Memory subtest (Woodcock, McGre, & Mather, 2001) and the Visual Working Memory subtest of the computerized *Colorado Assessment Tests 1.2* (Davis and Keller, 2002), respectively. For AWM, 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” followed by “4, 7, 2”). Age-normed standard scores were used for all statistical analyses. For VWM, participants were instructed to monitor a computer screen displaying eight 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.

## 2.3. Neural encoding of speech in noise

### 2.3.1. Stimulus and conditions

The evoking stimulus was a six-formant, 170 ms speech syllable /da/ that was synthesized using a Klatt-based synthesizer (Klatt, 1980) with a 5 ms voice onset time and a level fundamental frequency (100 Hz). The first, second and third formants were dynamic over the first 50 ms ( $F_1$ , 400–720 Hz;  $F_2$ , 1700–1240 Hz;  $F_3$ ,

2580–2500 Hz) and then maintained frequency for the remainder of the syllable. The fourth, fifth and sixth formants were constant throughout the duration of the stimulus ( $F_4$ , 3300 Hz;  $F_5$ , 3750 Hz;  $F_6$ , 4900 Hz). The stimulus was presented with an 81 ms inter-stimulus interval using NeuroScan Stim2 (Compumedics; Charlotte, NC, USA). For the noise condition, the stimulus was presented in the midst of a background of multi-talker babble. The 45 s noise file was created through the superimposition of grammatically correct but semantically anomalous sentences spoken by six different speakers (two males and four females) in Cool Edit Pro, version 2.1 (Syntrillium Software, Corp.; Scottsdale, AZ). This noise file employed recorded sentences that were originally designed for a study published by Smiljanic and Bradlow (2005). The SNR was set at +10 dB (da/noise) based on the root mean square (RMS) amplitude of the entire noise track.

### 2.3.2. Recording parameters

Six-thousand artifact-free auditory brainstem responses were recorded to the speech sound /da/ in NeuroScan Acquire 4.3 (Compumedics) using Ag–AgCl electrodes that were arranged in a vertical montage, with active at Cz (top of the head), ground at FPz (middle of the forehead) and the right earlobe as reference. Artifact rejection was monitored online and maintained at <10% (i.e., less than 6600 trials were presented to any given subject). Contact impedance for all electrodes was under 5 k $\Omega$ , with less than 3 k $\Omega$  difference across electrodes. The evoking stimulus was presented in 23–25 min blocks (quiet then noise) to the right ear in alternating polarities at 80 dB SPL through insert earphones (ER-3; Etymotic Research, Inc., Elk Grove Village, IL). During the recording, participants were seated in a sound-attenuated booth and watched movies of their choice, with soundtracks playing at  $\leq$ 40 dB. This method has proven successful for minimizing myogenic activity and maintaining subject alertness (Skoe & Kraus, 2010).

### 2.3.3. Data processing and analysis

Continuous neural recordings for quiet and noise conditions were off-line filtered from 70 to 2000 Hz (12 dB/octave, zero-phase shift) in NeuroScan Edit to minimize low-frequency myogenic noise and cortical activity and to include energy that would be expected in the brainstem response given its phase-locking limits (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010), epoched from –40 to 190 ms referenced to the presentation of the stimulus (0 ms), and baseline corrected. Responses with amplitudes  $>+/-35 \mu\text{V}$  were rejected as artifact and for each stimulus polarity 3000 artifact-free responses were averaged together and subsequently added. All data processing was executed with scripts generated in MATLAB 7.5.0 (The Mathworks, Natick, MA).

The SNR of the final average response was measured by dividing the RMS of the response (0–190 ms) by the RMS of the prestimulus period (–40 to 0 ms). This metric was used to ensure that the response was adequately free of myogenic and electrical noise; all participants demonstrated SNRs  $>1.5 \mu\text{V}$  in the quiet condition. Furthermore, musician and nonmusician groups were not distinct with regard to quiet response SNRs ( $F_{(1,29)} = 0.17$ ,  $p = 0.69$ ).

### 2.3.4. Auditory brainstem response timing

To gauge the effects of noise and musicianship on the timing of the neural response, three response peaks corresponding to the onset of the neural response (mean<sub>QUIET</sub> = 9.08,  $SD = 0.48$ ), to the formant transition of the stimulus (mean<sub>QUIET</sub> = 43.80,  $SD = 0.30$ ) and to the onset of the sustained vowel region (mean<sub>QUIET</sub> = 63.01,  $SD = 0.18$ ) (Fig. 3A) were identified and their latencies were compared across quiet and noise conditions. Peaks were first identified by a rater who was blind to the participants’ group characteristics and they were subsequently confirmed by the primary author. In the case of disagreement with peak identification,

the advice of a third peak picker, also blind to participants' group characteristics, was sought. All peaks were clearly identifiable in all participants.

### 2.3.5. Stimulus-to-response fidelity

To quantify the effect of noise on the fidelity of the neural response to the steady-state vowel (i.e., the section of the response that best resembles the waveform of the evoking stimulus), the stimulus and response waveforms were compared via cross-correlation. The degree of similarity was calculated by shifting the vowel section of the stimulus (50–170 ms) over a 7–12 ms range relative to the response, until a maximum correlation was found between the vowel portion of the stimulus and the corresponding steady-state response. This time lag (7–12 ms) was chosen because it encompasses the stimulus transmission delay (from the ER-3 transducer and ear insert  $\sim 1.1$  ms) and the neural lag between the cochlea and the rostral brainstem. This calculation resulted in Pearson's  $r$  values for both the quiet and noise conditions, which were Fisher transformed for all statistical analyses.

### 2.3.6. Spectral encoding

To assess the neural encoding of the stimulus spectrum, we applied a fast Fourier transform to the steady-state portion of the response (60–180 ms). From the resulting amplitude spectrum, average spectral amplitudes of specific frequency bins were calculated. Each bin was 20 Hz wide and centered on the fundamental frequency of the stimulus ( $f_0$ : 100 Hz) and the subsequent harmonics  $H_2$ – $H_8$  (200–800 Hz).

### 2.3.7. Statistical analysis

Musician and nonmusician behavioral data were compared using a one-way ANOVA. To gauge auditory brainstem response degradation, response measures in quiet and noise were compared using a Repeated Measures ANOVA (RMANOVA) with condition (quiet/noise) as within-subject variable and group (Mus/NonMus) as between-subject variable. Independent samples  $t$ -tests were employed to better define the effects observed. Relationships among musical practice histories and behavioral and neural data were examined with Pearson's correlations, partialling for IQ when appropriate. All results reported herein reflect two-tailed values and normality for all data was confirmed using the Kolmogorov–Smirnov test for equality. Statistics were computed using SPSS (SPSS, Inc., Chicago, IL).

## 3. Results

### 3.1. Summary of results

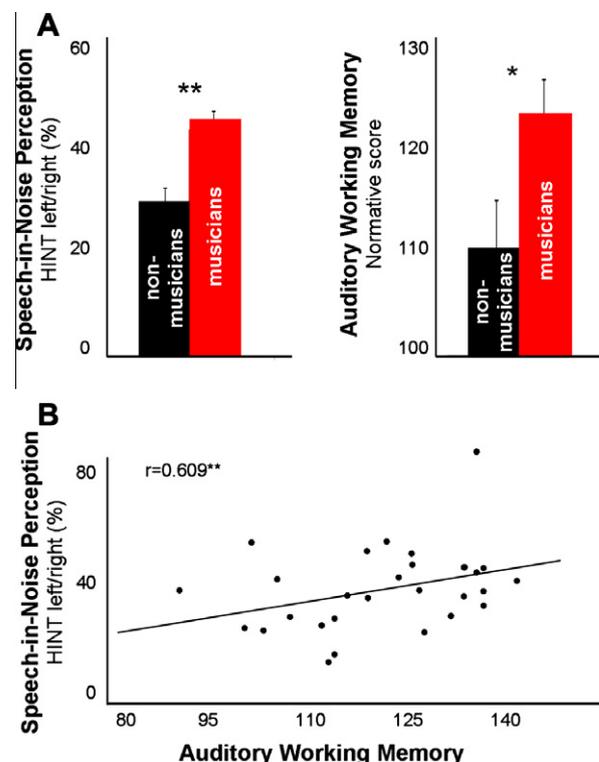
Musically trained children outperformed nonmusicians on speech-in-noise perception when the two signals were spatially segregated (HINT<sub>right</sub> and HINT<sub>left</sub>), as well as on the auditory working memory and auditory and visual attention tasks. Musicians also demonstrated less auditory brainstem response degradation with the addition of background noise compared to nonmusicians, although neural enhancements in musicians were observed in both quiet and noise conditions. Perceptual (speech in noise) and cognitive (auditory working memory, attention) performance correlated with auditory brainstem function as well as with musicians' extent of musical training.

### 3.2. Children with musical training have enhanced speech-in-noise, attention and auditory working memory abilities

Musically trained children outperformed nonmusicians on SIN perception for the higher-context assessment when the speech

and noise signals were spatially separated (Fig. 1A, left panel) (2 group<sub>mus/non</sub>  $\times$  3 noise condition<sub>front/left/right</sub> RMANOVA, main effect of condition:  $F_{(1,29)} = 8.28$ ,  $p = 0.002$ , group  $\times$  condition interaction:  $F_{(1,29)} = 2.70$ ,  $p = 0.08$ ; post hoc one-way ANOVAs, HINT<sub>left/right</sub>:  $F_{(1,29)} = 7.14$ ,  $p < 0.01$ , HINT<sub>front</sub>:  $F_{(1,29)} = 0.03$ ,  $p = 0.87$ ). No group differences were observed for SIN performance on the assessment that more directly reflects peripheral hearing ability (WIN:  $F_{(1,29)} = 2.39$ ,  $p = 0.14$ ) (Wilson et al., 2007). Musicians also demonstrated heightened auditory but not visual working memory (Fig. 1B, right panel) (2 group<sub>mus/non</sub>  $\times$  2 condition<sub>aud/vis</sub> RMANOVA, main effect of condition:  $F_{(1,29)} > 100$ ,  $p < 0.0001$ , group  $\times$  condition interaction:  $F_{(1,29)} = 7.42$ ,  $p = 0.01$ ; post hoc one-way ANOVAs, AWM:  $F_{(1,29)} = 4.87$ ,  $p < 0.05$ , VWM:  $F_{(1,29)} = 0.01$ ,  $p = 0.92$ ). With regard to attention, we were unable to collect attention data in two musicians due to time constraints; data in four nonmusicians were omitted from analyses due to noncompliance that resulted in excessively variable performance (scores could not be computed by the program). Although musicians and nonmusicians only marginally differed on the auditory attention quotient (2 group<sub>mus/non</sub>  $\times$  2 condition<sub>aud/vis</sub> RMANOVA, no main effect of condition:  $F_{(1,23)} < 1.0$ ,  $p > 0.5$  or group  $\times$  condition interaction:  $F_{(1,23)} = 1.91$ ,  $p = 0.18$ ; one-way ANOVAs, AAtt:  $F_{(1,23)} =$ ,  $p = 0.06$ , VISATT:  $F_{(1,23)} =$ ,  $p = 0.11$ ), the comparative power was reduced due to a smaller  $N$ .

Given co-variance among IQ (WASI), AWM and HINT performance (Table 1), we explored correlations between AWM and HINT performance with IQ treated as a covariant. Because IQ did not correlate with WIN and attention performance, it was not covaried in AWM-WIN and attention-HINT/attention-WIN comparisons. As stated previously, musicians and nonmusicians did not differ according to IQ. HINT performance and AWM correlated across all participants, with higher HINT<sub>left/right</sub> perception relating with better AWM (Table 1; Fig. 1B). Relationships were not observed between WIN and AWM performance, nor between



**Fig. 1.** Musicians demonstrate better speech perception in noise and auditory working memory than nonmusicians. Musicians' increased auditory working memory capacity may contribute to their enhanced speech-in-noise perception. \* $p < 0.05$ , \*\* $p < 0.01$ .

**Table 1**

Relationships among cognitive abilities and speech-in-noise perception as measured by the Words in Noise (WIN) and Hearing in Noise (HINT) Tests. IQ is held constant where appropriate (see Section 3).

Pearson's <i>r</i> values	WIN	HINT	IQ
IQ	0.06	0.33 <sup>~</sup>	–
Auditory working memory	0.18	0.61 <sup>**</sup>	0.51 <sup>**</sup>
Visual working memory	0.004	0.10	0.32
Auditory attention	0.02	0.24	0.13
Visual attention	0.21	0.30	0.30

<sup>~</sup>*p* < 0.05.

<sup>~</sup> *p* < 0.1.

<sup>\*\*</sup> *p* < 0.01.

either measure of speech-in-noise perception and auditory or visual attention performance.

### 3.3. Neural encoding of speech in noise

#### 3.3.1. Stimulus-to-response fidelity

Musicians demonstrated less quiet-to-noise auditory brainstem response degradation compared to nonmusicians over the vowel portion of the response, as revealed by stimulus-to-response correlations (Fig. 2B, left panel). Specifically, we observed an interaction between group<sub>mus/non</sub> and condition<sub>quiet/noise</sub> (Fig. 2C) ( $2 \times 2$  RMANOVA:  $F_{(1,29)} = 29.00$ ,  $p < 0.02$ ). An independent samples *t*-test considering the degree of stimulus-to-response correlation decrease from quiet to noise ( $Z_{\text{quiet}} - Z_{\text{noise}}$ ) revealed that musicians have less response degradation with the addition of noise, relative to nonmusicians ( $t_{(30)} = 3.03$ ,  $p = 0.005$ ). Although musicians and nonmusicians had equivalent stimulus-to-response correlations in the quiet condition ( $t_{(30)} = 1.14$ ,  $p = 0.27$ ), musicians had higher stimulus-to-response correlations in the noise condition (Fig. 2C;  $t_{(30)} = 4.48$ ,  $p < 0.01$ ), reflecting maintained stimulus-to-response fidelity in the presence of background noise.

#### 3.3.2. Auditory brainstem response timing

We compared response timing within conditions as well as the quiet-to-noise timing shifts for the onset, transition and sustained response peaks (occurring at ~9, 43 and 63 ms, respectively). A  $2$  (group)  $\times 2$  (condition)  $\times 3$  (peak) RMANOVA revealed a significant main effects of group and noise on auditory brainstem response timing, with responses in noise occurring later than responses in quiet ( $F_{(1,29)} = 53.87$ ,  $p < 0.0001$ ) but with musicians' responses occurring earlier than nonmusicians' ( $F_{(1,29)} = 10.72$ ,  $p < 0.005$ ). Independent samples *t*-tests indicate that musicians' response peaks within the transition region in both quiet and noise conditions were earlier than nonmusicians' (peak 43<sub>QUIET</sub>:  $t_{(30)} = 2.65$ ,  $p < 0.01$ ; peak 43<sub>NOISE</sub>:  $t_{(30)} = 2.62$ ,  $p < 0.01$ ), without significant group differences among the onset and vowel region peaks (all  $t < 1.6$ ,  $p > 0.1$ ). This indicates that musicians demonstrated faster responses than nonmusicians only to the most spectrotemporally-dynamic and informationally-salient portion of the stimulus, the formant transition.

With regard to quiet-to-noise timing shifts, there was a significant interaction between noise and musician grouping ( $F_{(1,29)} = 4.72$ ,  $p < 0.05$ ). Independent samples *t*-test revealed that in their responses to noise relative to their responses in quiet, musicians demonstrated less of a delay in the formant transition region than nonmusicians (peak 43; Fig. 3B and C) ( $t_{(30)} = 2.00$ ,  $p < 0.05$ ). Group differences for the timing shift of the other two peaks did not approach significance (all  $t < 1.3$ ,  $p > 0.2$ ). The relevance of auditory brainstem response timing in the encoding of the formant transition of this same speech stimulus for speech-in-noise processing has previously been established (Anderson et al., 2010; Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009; Parbery-Clark et al., 2009).

#### 3.3.3. Spectral encoding

In both quiet and noise conditions, musicians demonstrated more robust auditory brainstem representation of the harmonics of the speech stimulus than nonmusicians (Fig. 2B, right panel). A  $2$  group<sub>mus/non</sub>  $\times 2$  condition<sub>quiet/noise</sub>  $\times 7$  harmonic<sub>H2-H8</sub> RMANOVA revealed main effects of noise ( $F_{(1,29)} = 93.83$ ,  $p < 0.0001$ ) and group ( $F_{(1,29)} = 8.62$ ,  $p < 0.01$ ), but no group  $\times$  condition interaction ( $F_{(1,29)} < 0.001$ ,  $p = 0.999$ ). No group differences were observed for the neural encoding of the fundamental frequency, within or across conditions.

### 3.4. Neural encoding of speech in noise relates to perceptual and cognitive performance

Auditory brainstem response characteristics correlated with speech-in-noise perception as well as with auditory, but not visual, working memory and attention (Table 2). Relationships were specifically apparent for the spectral encoding of the speech stimulus, in which the averaged magnitude of the second through eighth harmonics in response to the speech stimulus in both quiet and noise conditions correlated with SIN perception (IQ partialled). Additionally, AWM correlated with the magnitude of H<sub>2</sub>–H<sub>8</sub> in both conditions (Fig. 4), unlike visual working memory (IQ partialled). Auditory and visual attention correlated with the magnitude of H<sub>2</sub>–H<sub>8</sub> in noise only (Fig. 4). With regard to neural response timing, the degree of quiet-to-noise shift (noise–quiet peak latencies, in ms) of the peak occurring at 43 ms in the transition region (Fig. 3B) correlated with auditory but not visual working memory (Fig. 4; IQ partialled) as well as with both auditory and, marginally, visual attention (Fig. 4).

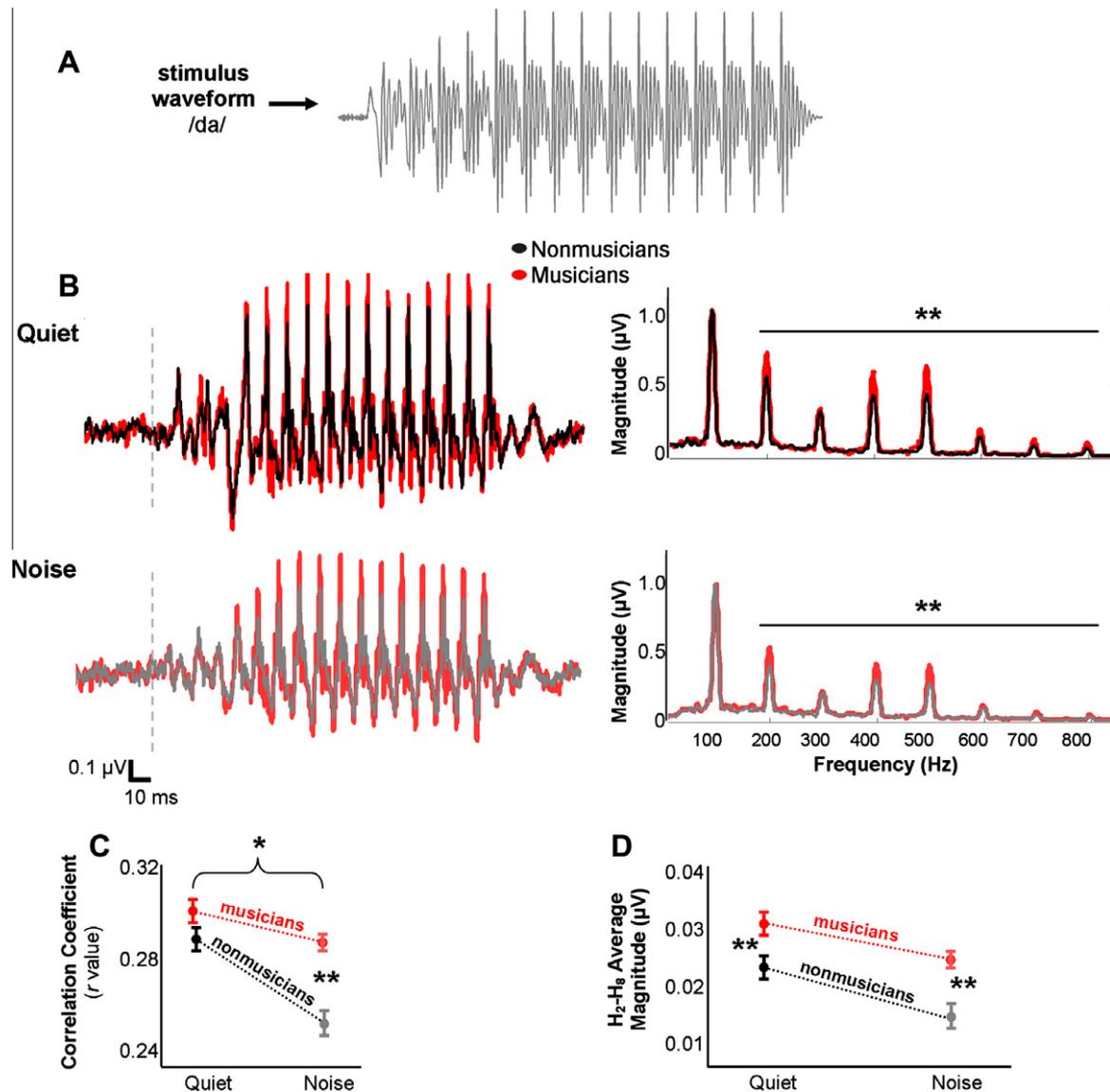
### 3.5. Relationships among speech-in-noise perception, auditory working memory, attention and extent of musical training

Perceptual, auditory cognitive (i.e., auditory but not visual working memory and attention) and spectral aspects of neural function (i.e., the encoding of speech harmonics) correlated with participants' extent of musical practice in that children with more years of musical training performed better and demonstrated more robust neural encoding of speech. These correlations were observed over all children with any degree of musical training, including three nonmusicians with a small amount of musical training, and considered years spent in pre-school music classes ( $N = 18$ ; HINT<sub>left/right</sub>:  $r = 0.629$ ,  $p < 0.001$ ; AAtt:  $r = 0.471$ ,  $p < 0.05$ ; VAtt:  $r = 0.198$ ,  $p = 0.43$ ; H<sub>2</sub>–H<sub>8</sub> QUIET:  $r = 0.461$ ,  $p < 0.03$ ; H<sub>2</sub>–H<sub>8</sub> NOISE:  $r = 0.562$ ,  $p < 0.005$ ). Although the correlation between years of musical training and AWM was only marginally significant per two-tailed values (AWM:  $r = 0.375$ ,  $p = 0.08$ ), VWM did not approach significance (VWM:  $r = -0.200$ ,  $p = 0.39$ ).

Although it could be argued that relationships to years of practice were driven by age, age did not correlate with either years of musical practice (considering only subjects with musical training histories:  $r = 0.22$ ,  $p = 0.31$ ) or auditory-specific cognitive abilities (HINT<sub>left/right</sub>:  $r = 0.29$ ,  $p = 0.12$ ; AAtt:  $r = 0.18$ ,  $p = 0.38$ ; AWM:  $r = 0.14$ ,  $p = 0.45$ ). Although the neural encoding of speech harmonics marginally related to age (H<sub>2</sub>–H<sub>8</sub> QUIET:  $r = 0.332$ ,  $p = 0.07$ ; H<sub>2</sub>–H<sub>8</sub> NOISE:  $r = 0.350$ ,  $p = 0.05$ ), the relationship between years of musical practice and this neural measure retained significance with age held constant (H<sub>2</sub>–H<sub>8</sub> QUIET:  $r = 0.385$ ,  $p < 0.05$ ; H<sub>2</sub>–H<sub>8</sub> NOISE:  $r = 0.378$ ,  $p < 0.05$ ).

## 4. Discussion

We reveal that the musician enhancement for the perception and neural encoding of speech in noise arises early in life, with



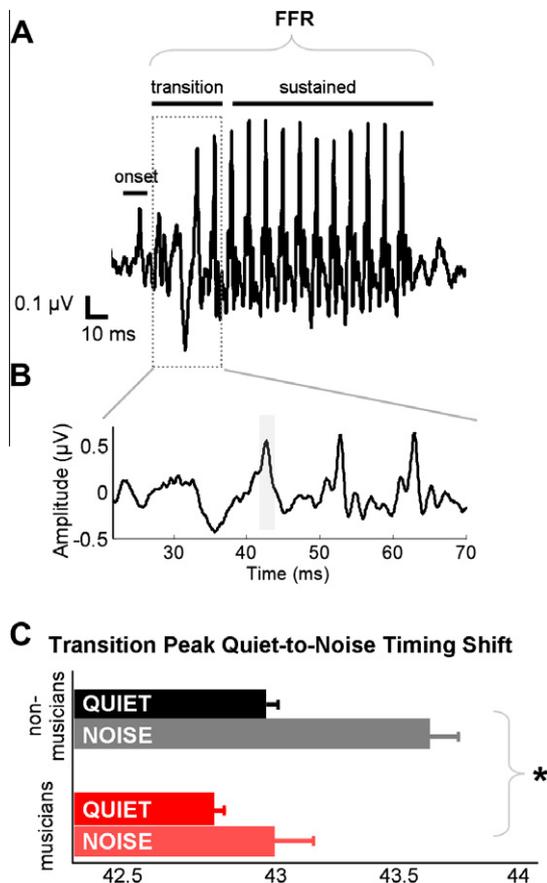
**Fig. 2.** Musicians' neural responses to speech are more resistant to the degradative effects of background noise than nonmusicians. Because of the phase-locking characteristics of auditory brainstem nuclei, auditory brainstem response waveforms (B) spectrally and temporally resemble the acoustic waveforms of evoking stimuli (A). In fact, the two signals can be cross-correlated to provide an index of neural stimulus-to-response fidelity. Although musicians and nonmusicians have equivalent neural fidelity in responses to speech in a quiet background, musicians experience less neural degradation with the addition of noise (C). Musicians also demonstrate more robust encoding of speech harmonics in both quiet and noise than nonmusicians (B, right panel; D). \* $p < 0.05$ , \*\* $p < 0.01$ .

more years of training relating with more robust speech processing in children. These perceptual and neural enhancements may be driven by the strengthening of auditory-specific cognitive abilities, such as auditory working memory and auditory attention, with musical training. Musicians and nonmusicians did not differ on tests of visual working memory and attention. Here, we discuss plausible mechanisms that may underlie these between-group differences from the perspective of co-existing innate differences between musicians and nonmusicians and neuroplastic changes with musical training, their developmental implications and potential contributions to clinical and educational efforts.

#### 4.1. Mechanisms of subcortical training-related plasticity and developmental implications

Although causation cannot be inferred from correlation, the relationships we report between years of musical practice and auditory brainstem function contribute to a growing literature supporting the modulation of auditory brainstem function with

interactive music (Bidelman, Gandour, & Krishnan, 2009; Bidelman & Krishnan, 2010; Musacchia et al., 2007; Parbery-Clark, Anderson, Hittner, & Kraus, 2012; Parbery-Clark, Tierney, Strait, & Kraus, 2012; Parbery-Clark et al., 2009; Strait, Kraus, Skoe, & Ashley, 2009; Wong et al., 2007) and language experience (Bidelman et al., 2009; Carcagno & Plack, 2011; Krishnan, Gandour, Bidelman, & Swaminathan, 2009; Krishnan, Swaminathan, & Gandour, 2008; Krishnan, Xu, Gandour, & Cariani, 2005; Song, Skoe, Banai, & Kraus, 2012; Song, Skoe, Wong, & Kraus, 2008). Whereas previous studies have demonstrated some degree of subcortical neuroplasticity in children with short-term software-based auditory training (Russo, Hornickel, Nicol, Zecker, & Kraus, 2010; Russo, Nicol, Zecker, Hayes, & Kraus, 2005), our findings present the first evidence that the non-experimental implementation of auditory training during early developmental years relates to, and may fundamentally shape, primary neurosensory function. An alternate but not mutually exclusive hypothesis is that children who begin musical training at a younger age (and thus would have received more years of practice) are genetically predisposed to have more robust auditory



**Fig. 3.** Musicians have faster neural timing to speech in quiet and noise. The speech stimulus and corresponding brainstem responses can be divided into regions that correspond to the spectrotemporally dynamic formant transition and the more periodic sustained vowel (A). Musicians have faster neural timing over the formant transition region than nonmusicians in both quiet and noise, in addition to less of a timing shift with the addition of background noise (C). \* $p < 0.05$ .

**Table 2**

Relationships among auditory brainstem response characteristics, cognitive abilities and speech-in-noise perception as measured by the Words in Noise (WIN) and Hearing in Noise (HINT) Tests (Pearson's  $r$  values). Auditory brainstem response characteristics include the magnitude of the response to the second through eighth harmonics and the response delay in the presence of noise. IQ is held constant where appropriate (see Section 3).

Pearson's $r$ values	H <sub>2</sub> –H <sub>8</sub> QUIET	H <sub>2</sub> –H <sub>8</sub> NOISE	Timing delay
IQ	0.33 <sup>~</sup>	0.34 <sup>~</sup>	–0.01
WIN (thresholds)	–0.38 <sup>*</sup>	–0.43 <sup>*</sup>	0.25
HINT (percentile ranks)	0.52 <sup>**</sup>	0.51 <sup>**</sup>	–0.25
Auditory working memory	0.55 <sup>**</sup>	0.51 <sup>**</sup>	–0.47 <sup>**</sup>
Visual working memory	0.02	–0.03	–0.06
Auditory attention	0.20	0.50 <sup>**</sup>	–0.45 <sup>~</sup>
Visual attention	0.14	0.46 <sup>*</sup>	–0.38 <sup>~</sup>

<sup>~</sup>  $p < 0.1$ .

<sup>\*</sup>  $p < 0.05$ .

<sup>\*\*</sup>  $p < 0.01$ .

brainstem function. This interpretation may be supported by work demonstrating relationships between music aptitude, language ability and auditory cortical function in nonmusician children, regardless of musical training (Milovanov, Huotilainen, Valimäki, Esquef, & Tervaniemi, 2008; Milovanov et al., 2009).

Consistent with known mechanisms of plasticity in auditory brainstem nuclei, we suggest that training-related changes reflect top-down neuromodulation of auditory brainstem response properties by cortical and thalamic activity. This interpretation could

account for the correlations reported here between auditory cognitive abilities and subcortical function, which may reflect strengthened cognitive control over basic sensory processing. There is no doubt that the auditory cortex is structurally and functionally shaped through interactive experience with sound, with changes being especially pronounced in responses to sounds with behavioral relevance (Fritz, Elhilali, & Shamma, 2007; Recanzone, Schreiner, & Merzenich, 1993). This plasticity is mediated by cholinergic inputs to the primary auditory cortex from the nucleus basalis (Kilgard & Merzenich, 1998; Metherate & Ashe, 1993), a region in the basal forebrain that is highly innervated by limbic nuclei involved in learning and memory. Animal models reveal that descending innervations from the auditory cortex can modulate auditory brainstem response characteristics (Ma & Suga, 2001) and that when these innervations are disabled, these changes do not occur (Bajo, Nodal, Moore, & King, 2010). Activation of the primary auditory cortex, especially when paired with behavioral reward (Suga & Ma, 2003), induces changes in subcortical response properties related to refined frequency and duration tuning and earlier response timing. Although local mechanisms of subcortical plasticity surely exist, the role that the descending auditory system plays to shape subcortical response properties is significant. Given that musical training relies on learning to associate slight acoustic discrepancies with behavioral significance and is dependent on memory, attention and emotional engagement, musical training may provide a remarkable avenue for inducing auditory plasticity in humans, especially during developmental years.

Faster auditory brainstem response timing and decreased response delays with noise in musically trained children are significant given the relationships these subcortical measures have with language-based developmental disorders (e.g., dyslexia, speech-in-noise impairment). Language-impaired children encode speech in quiet similarly to typically developing children but demonstrate increased subcortical delays with the addition of noise (specifically, at 43 ms in response to our same speech stimulus; Fig. 3B), indicating compromised neural function in adverse listening environments (Anderson et al., 2010). The diversity of our everyday listening environments necessitates neuronal adaptability for speech encoding, permitting the nervous system not only to meet the sensory demands of quiet and noisy circumstances but also to overcome different levels and types of noise. Here, we suggest that musical training during early childhood reduces the delays imposed by noise in the auditory brainstem response to speech. This outcome may indicate that musical training engenders more adaptable nervous systems.

In addition to faster neural timing, musician children also demonstrate more robust encoding of speech harmonics in quiet and noise. This observation again relates to what has been observed in reading- and language-impaired children, who demonstrate decreased encoding of speech harmonics in quiet (Anderson, Skoe, Chandrasekaran, Zecker, & Kraus, 2010; Banai et al., 2009). In fact, the magnitude of the neural response to speech harmonics correlates with reading ability (Banai et al., 2009) and SIN perception (Anderson et al., 2010). The relevance of harmonics for speech-in-noise perception likely relies on how speech harmonics guide the perceptual differentiation of voices, enabling us to distinguish speakers from another based on their voice quality – especially in the presence of competing voices.

#### 4.2. The auditory brainstem response reflects cognitive contributors to sensory function

Here, we demonstrate that auditory attention and working memory correlate with auditory brainstem response properties. These outcomes are supported by recent work revealing a relationship between executive function and the subcortical

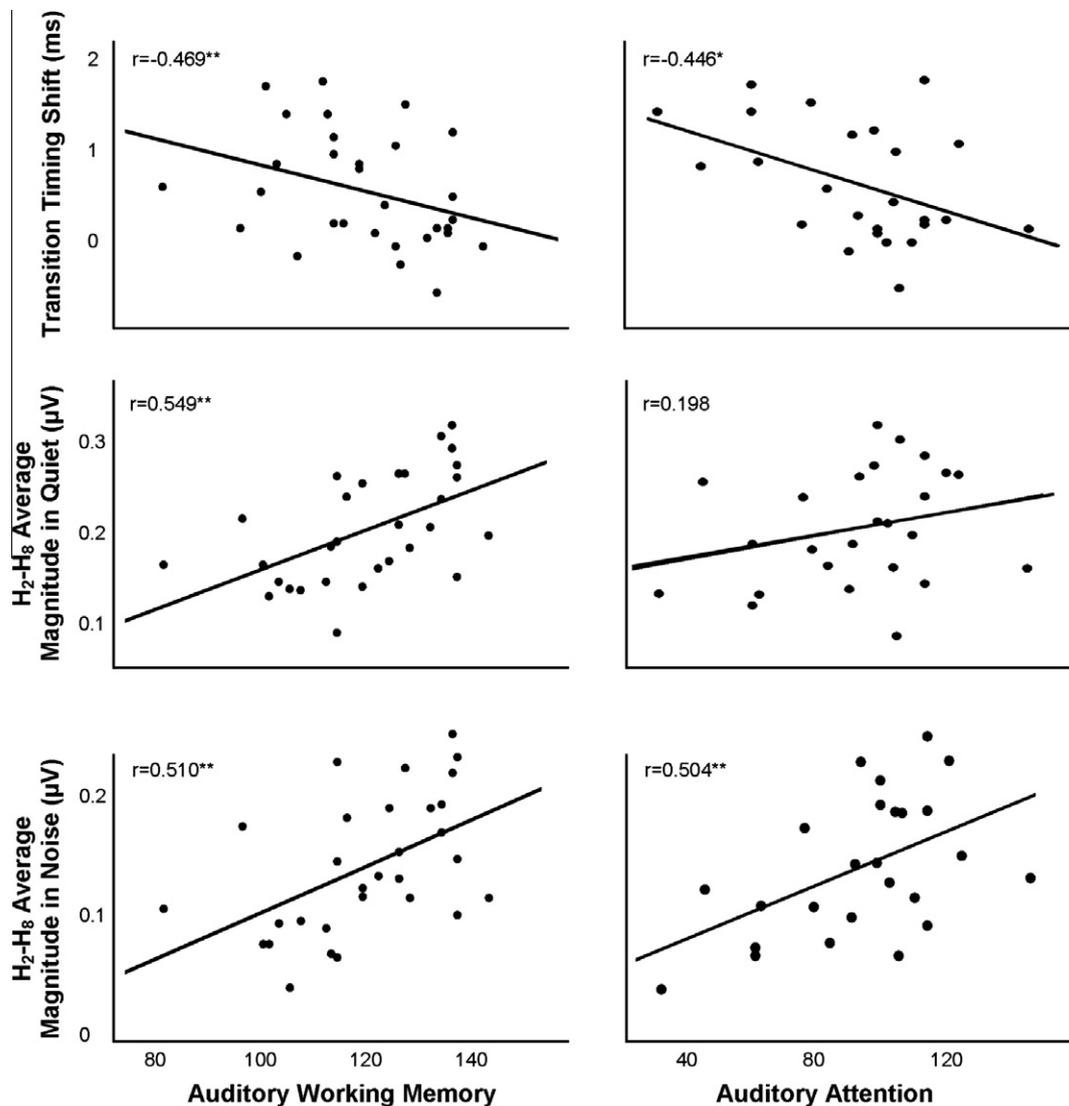


Fig. 4. Auditory working memory and attention correlate with the neural encoding of speech in quiet and noise. \* $p < 0.05$ , \*\* $p < 0.01$ .

processing of pitch in bilinguals (Krizman, Marian, Shook, Skoe, & Kraus, 2012). Although we cannot test our hypothesis in the present study, we propose that these relationships reflect top-down cognitive contributions to auditory brainstem function.

Cognitive contributors to primary sensory function in auditory cortex have been well-established (Fritz, Elhilali, David, & Shamma, 2007; Fritz et al., 2007; Weinberger, 2004). This work is strengthened by functional connectivity between extra-sensory cortices associated with executive function (e.g., prefrontal cortex, anterior cingulate cortex) and primary auditory cortex (Crottaz-Herbette & Menon, 2006; Fritz, David, Radtke-Schuller, Yin, & Shamma, 2010; Morris, Friston, & Dolan, 1998; Pandya, Van Hoesen, & Mesulam, 1981). Given that the mammalian auditory system houses extensive descending corticocollicular and corticothalamic tracts, which are even thought to surpass the volume of ascending fibers (He, 2003), top-down cognitive contributions to auditory brainstem response properties are not only possible, but likely. Future work in animal models might assess contributions of extra-sensory cortices to subcortical response properties; observations over the course of learning may reveal strengthened task-specific cognitive-subcortical connectivity with increased training, especially during sensitive developmental periods.

#### 4.3. Developmental implications

Although previous work in adults has hinted at developmental benefits of musical training for speech perception and neural encoding in noise, our findings provide the first evidence that musical training relates with improved speech-in-noise processing during childhood. Parbery-Clark and colleagues have demonstrated enhanced speech-in-noise perception in adult musicians (Parbery-Clark et al., 2009, 2011) for the most perceptually difficult condition (HINT<sub>front</sub>, for which the speech and noise are collocated) and not when the two signals are spatially segregated. Musician children, on the other hand, outperform nonmusicians on the spatially segregated task but not when the speech and noise are collocated. These results can be interpreted within a developmental framework by considering that the collocated condition is more perceptually challenging than the spatially-separated condition, in which the speech signal can be segregated based on both acoustic and spatial cues. We propose that musical training accelerates the development of speech-in-noise perception by first impacting performance on this easier listening condition.

With regard to brain function, child musicians demonstrate enhanced neural timing and harmonic representation of speech in both quiet and noise; adult musicians only demonstrate these

enhancements in noise, the more challenging of the two conditions. We propose that, even amidst innate differences that may distinguish children who undergo musical training from those who do not, musical training during development steepens the neurodevelopmental trajectory of these mechanisms irrespective of the listening environment. Although the musician advantage in quiet fades, musicians' faster neural timing and more robust encoding of speech harmonics in the presence of noise persists into adulthood. More work, especially longitudinal work, must be pursued to confirm our interpretation and to elucidate the developmental time-course of perceptual and cognitive enhancements and neural plasticity with musical training.

#### 4.4. Addressing contributions of nature and nurture

It is often asked: are differences between musically trained and untrained children solely the outcome of genetic predispositions? The correlations reported here between musical training histories and neural and perceptual indices of speech-in-noise processing point to non-genetic determinants, at least in part. Auditory-specific cognitive performance (i.e., auditory but not visual working memory and attention) correlated with the number of years that children had undergone musical training; these relationships have also been reported in musician adults (Parbery-Clark et al., 2009; Strait & Kraus, 2011a; Strait et al., 2010). The neural benefits of music training relate not only to years of practice in musicians but also to years of *past* music experience, observed in nonmusician adults following many years of discontinued practice (Skoe & Kraus, 2012). Although it is plausible that the parents of children with weaker auditory skills do not invest in music lessons or that children with stronger auditory processing skills persevere with music training, we argue that musical training plays a role in strengthening the neural and cognitive underpinnings of speech perception.

It is clear that the potential for success as a musician is not ubiquitous. Rather, children who persist with consistent musical training into adulthood comprise a small percentage of the musically trained population; their artistry appears to be underpinned, at least in part, by a complex genetic tapestry (Ukkola, Onkamo, Raijas, Karma, & Jarvela, 2009). Here, we contribute to the literature concerning innate and training-related aspects of musician/nonmusician distinctions by proposing that neither genetic contributors nor early predisposition toward musical skill can fully predict the pursuit of musical training an individual level; rather, a diverse range of socio-cultural factors promote one's pursuit and continuation of musical training (Burland & Davidson, 2002; Davidson, Howe, Moore, & Sloboda, 1998; Moore, Burland, & Davidson, 2003). Providing access to musical training during early childhood years may be of developmental importance for all children by facilitating the strengthening of neural mechanisms that underlie auditory perceptual and cognitive performance. Further studies should assess the impact of music on speech-evoked auditory brainstem activity and auditory perceptual and cognitive performance using longitudinal designs.

#### 4.5. Conclusions

Taken together, we reveal benefits in musicians for hearing in noise during pivotal developmental years, with child musicians demonstrating strengthened neural encoding of key acoustic ingredients for speech perception in challenging listening environments compared to nonmusicians. Musicians' auditory processing enhancements may be driven, at least in part, by strengthened auditory cognitive-function with musical training. Given that musicians demonstrate speech-in-noise processing enhancements for the very aspects of neural function that underlie language skills, these results bear relevance for educators, scientists and clinicians

involved in the assessment and remediation of language-based learning deficits.

#### Acknowledgments

The authors wish to thank S. O'Connell for her assistance with data management and V. Abecassis, K. Chan and A. Choi for their assistance with data collection. This work was supported by NIH F31DC011457-01 to D.S., NSF 0921275 to N.K. and by the Grammy Research Foundation.

#### References

- Anderson, S., Chandrasekaran, B., Skoe, E., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *Journal of Neuroscience*, 30(14), 4922–4926.
- Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010). Brainstem correlates of speech-in-noise perception in children. *Hearing Research*, 270, 151–157.
- Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, 13(2), 253–260.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex*, 19(11), 2699–2707.
- Bangert, M., & Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *European Journal of Neuroscience*, 24(6), 1832–1834.
- Besson, M., Schon, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, 25(3–4), 399–410.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2009). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, 23(2), 425–434.
- Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, 1355, 112–125.
- Burland, K., & Davidson, J. W. (2002). Training the talented. *Music Education Research*, 4, 121–140.
- Carcagno, S., & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, 12(1), 89–100.
- Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature*, 396(6707), 128.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, 47, 236–246.
- Chobert, J., Marie, C., Francois, C., Schon, D., & Besson, M. (2011). *Enhanced passive and active processing of syllables in musician children*. *Journal of Cognitive Neuroscience*.
- Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review*, 8(2), 331–335.
- Crottaz-Herbette, S., & Menon, V. (2006). Where and when the anterior cingulate cortex modulates attentional response: Combined fMRI and ERP evidence. *Journal of Cognitive Neuroscience*, 18(5), 766–780.
- Davidson, J. W., Howe, M. J. A., Moore, D. M., & Sloboda, J. A. (1998). The role of teachers in the development of musical ability. *Journal of Research in Music Education*, 46, 141–160.
- Davis, H. P., & Keller, F. (2002). *Colorado assessment tests (CATS), version 1.2*. Colorado Springs, CO: Colorado Assessment Tests.
- Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *Journal of the Acoustical Society of America*, 66(3), 651–653.
- Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*, 108(6), 3023–3029.
- Flege, J. E., Meador, D., & MacKay, I. R. A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism*, 3(1), 55–67.
- Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P., & Shamma, S. A. (2010). Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nature Neuroscience*, 13(8), 1011–1019.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research*, 229(1–2), 186–203.
- Fritz, J. B., Elhilali, M., & Shamma, S. A. (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of Neurophysiology*, 98(4), 2337–2346.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, 129(10), 2593–2608.
- Gaab, N., & Schlaug, G. (2003). Musicians differ from nonmusicians in brain activation despite performance matching. *Annals of the New York Academy of Science*, 999, 385–388.

- Hall, J. W., 3rd, Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing, 23*(2), 159–165.
- He, J. (2003). Corticofugal modulation of the auditory thalamus. *Experimental Brain Research, 153*(4), 579–590.
- 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. *Proceedings of the National Academy of Sciences of the USA, 106*, 13022–13027.
- Hutchinson, S., Lee, L. H., Gaab, N., & Schlaug, G. (2003). Cerebellar volume of musicians. *Cerebral Cortex, 13*(9), 943–949.
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., et al. (2009). Musical training shapes structural brain development. *Journal of Neuroscience, 29*(10), 3019–3025.
- Jakobson, L. S., Lewycky, S. T., Kilgour, A. R., & Stoesz, B. M. (2008). Memory for verbal and visual material in highly trained musicians. *Music Perception, 26*(1), 41–55.
- Kilgard, M. P., & Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science, 279*(5357), 1714–1718.
- Klatt, D. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67*, 13–33.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605.
- Kraus, N., Parbery-Clark, A., & Strait, D. L. (2012). Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Annals of the New York Academy of Sciences: The Neurosciences and Music IV, 1252*, 100–107.
- Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *NeuroReport, 20*(4), 408–413.
- Krishnan, A., Swaminathan, J., & Gandour, J. T. (2008). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience, 21*(6), 1092–1105.
- Krishnan, A., Xu, Y., Gandour, J., & Ciani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research Cognitive Brain Research, 25*(1), 161–168.
- 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. *Proceedings of the National Academy of Sciences of the United States of America, 109*(20), 7877–7881.
- Ma, X., & Suga, N. (2001). Plasticity of bat's central auditory system evoked by focal electric stimulation of auditory and/or somatosensory cortices. *Journal of Neurophysiology, 85*(3), 1078–1087.
- Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199–211.
- Metherate, R., & Ashe, J. H. (1993). Nucleus basalis stimulation facilitates thalamocortical synaptic transmission in the rat auditory cortex. *Synapse, 14*(2), 132–143.
- Milovanov, R., Huotilainen, M., Esquef, P. A., Alku, P., Valimaki, V., & Tervaniemi, M. (2009). The role of musical aptitude and language skills in preattentive duration processing in school-aged children. *Neuroscience Letters, 460*(2), 161–165.
- Milovanov, R., Huotilainen, M., Valimaki, V., Esquef, P. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research, 1194*, 81–89.
- Moore, D. G., Burland, K., & Davidson, J. W. (2003). The social context of musical success: A developmental account. *British Journal of Psychology, 94*, 529–545.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*, 1203–1207.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*(3), 712–723.
- Morris, J. S., Friston, K. J., & Dolan, R. J. (1998). Experience-dependent modulation of tonotopic neural responses in human auditory cortex. *Proceedings of the Royal Society B: Biological Sciences, 265*(1397), 649–657.
- Muchnik, C., Ari-Even Roth, D., Othman-Jebar, R., Putter-Katz, H., Shabtai, E. L., & Hildesheimer, M. (2004). Reduced medial olivocochlear bundle system function in children with auditory processing disorders. *Audiology and Neurotology, 9*(2), 107–114.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America, 104*(40), 15894–15898.
- Nilsson, M., Soli, S., & Sullivan, J. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America, 95*, 1085–1099.
- Pandya, D. N., Van Hoesen, G. W., & Mesulam, M. M. (1981). Efferent connections of the cingulate gyrus in the rhesus monkey. *Experimental Brain Research, 42*(3–4), 319–330.
- Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging, 33*, 1483.e1–1483.e4.
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience, 29*(45), 14100–14107.
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*(6), 653–661.
- Parbery-Clark, A., Strait, D. L., 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*(5), e18082.
- Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*.
- Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience, 13*(1), 87–103.
- Russo, N. M., Hornickel, J., Nicol, T., Zecker, S., & Kraus, N. (2010). Biological changes in auditory function following training in children with autism spectrum disorders. *Behavioral and Brain Functions, 6*, 60.
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research, 156*(1), 95–103.
- Sandford, J. A., & Turner, A. (1994). *Intermediate visual and auditory continuous performance test (Version 2.6)*. Richmond, VA: Braintrain.
- Schlaug, G. (2001). The brain of musicians: A model for functional and structural adaptation. *Annals of the New York Academy of Science, 930*, 281–299.
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences, 1060*, 219–230.
- Seppanen, M., Brattico, E., & Tervaniemi, M. (2007). Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory, 87*(2), 236–247.
- Seppanen, M., Pesonen, A. K., & Tervaniemi, M. (2012). Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. *Attention, Perception, and Psychophysics, 74*(3), 600–612.
- Shahin, A., Roberts, L. E., & Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *NeuroReport, 15*(12), 1917–1921.
- Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing, 31*(3).
- Skoe, E., & Kraus, N. (2012). A little goes a long way: how the adult brain is shaped by musical training in childhood. *Journal of Neuroscience, 32*(34), 11507–11510.
- Smiljanic, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118*(3 Pt 1), 1677–1688.
- Song, J., Skoe, E., Banai, K., & Kraus, N. (2010). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*.
- Song, J., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex, 22*, 1180–1190.
- Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience, 20*(10), 1892–1902.
- Sperling, A. J., Lu, Z. L., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8*(7), 862–863.
- 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. *Developmental Science, 12*(4), 634–646.
- Strait, D. L., & Kraus, N. (2011a). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front Psychology, 2*, 113.
- Strait, D. L., & Kraus, N. (2011b). Playing music for a smarter ear: Cognitive, perceptual and neurobiological evidence. *Music Perception, 29*(2), 133–147.
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research, 261*, 22–29.
- Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency: Effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29*(3), 661–668.
- Suga, N., & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neuroscience, 4*(10), 783–794.
- Tallal, P., & Stark, R. E. (1981). Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America, 69*(2), 568–574.
- Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience, 30*(8), 1636–1642.
- Ukkola, L. T., Onkamo, P., Raijas, P., Karma, K., & Jarvela, I. (2009). Musical aptitude is associated with AVPR1A-haplotypes. *PLoS ONE, 4*(5), e5534.
- Van Tasell, D. J., Hagen, L. T., Koblas, L. L., & Penner, S. G. (1982). Perception of short-term spectral cues for stop consonant place by normal and hearing-impaired subjects. *Journal of the Acoustical Society of America, 72*, 1771–1780.
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience, 5*(4), 279–290.
- Wilson, R. H. (1993). Development and use of auditory compact discs in auditory evaluation. *Journal of Rehabilitation Research and Development, 30*(3), 342–351.
- Wilson, R. H., McArdle, R. A., & Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and

- listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*, 50(4), 844–856.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422.
- Woodcock, R. W., McGre, K. S., & Mather, N. (2001). *Woodcock–Johnson psycho-educational battery* (3rd ed.). Itasca, IL: Riverside.
- Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., & Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), 14110–14115.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12(5), 732–745.