

## Biological impact of preschool music classes on processing speech in noise



Dana L. Strait<sup>a,b,1</sup>, Alexandra Parbery-Clark<sup>b,c</sup>, Samantha O'Connell<sup>b</sup>,  
Nina Kraus<sup>a,b,c,d,e,\*</sup>

<sup>a</sup> Institute for Neuroscience, Northwestern University, Chicago, IL 60611, USA

<sup>b</sup> Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL 60208, USA

<sup>c</sup> Department of Communication Sciences, Northwestern University, Evanston, IL 60208, USA

<sup>d</sup> Department of Neurobiology & Physiology, Northwestern University, Evanston, IL 60208, USA

<sup>e</sup> Department of Otolaryngology, Northwestern University, Evanston, IL 60208, USA

### ARTICLE INFO

#### Article history:

Received 20 March 2013

Received in revised form 13 June 2013

Accepted 14 June 2013

#### Keywords:

Auditory  
ABR  
Brainstem  
Language  
Musicians  
Music training  
Subcortical  
Development

### ABSTRACT

Musicians have increased resilience to the effects of noise on speech perception and its neural underpinnings. We do not know, however, how early in life these enhancements arise. We compared auditory brainstem responses to speech in noise in 32 preschool children, half of whom were engaged in music training. Thirteen children returned for testing one year later, permitting the first longitudinal assessment of subcortical auditory function with music training. Results indicate emerging neural enhancements in musically trained preschoolers for processing speech in noise. Longitudinal outcomes reveal that children enrolled in music classes experience further increased neural resilience to background noise following one year of continued training compared to nonmusician peers. Together, these data reveal enhanced development of neural mechanisms undergirding speech-in-noise perception in preschoolers undergoing music training and may indicate a biological impact of music training on auditory function during early childhood.

© 2013 Elsevier Ltd. All rights reserved.

### 1. Introduction

The majority of learning during early childhood, including language learning, occurs in noisy acoustic environments (Bradley and Sato, 2008). Hearing speech in noise, however, is a poignant challenge for children, who require higher signal-to-noise ratios than adults for

comparable performance (Elliott, 1979; Fallon et al., 2000; Hall et al., 2002; Papso and Blood, 1989). Although hearing thresholds at birth are comparable to those of adults (Sininger et al., 1997), the maturational trajectories of perceptual abilities that undergird speech-in-noise perception likely contribute to its protracted development. For example, spatial discrimination (Litovsky, 1997), frequency resolution (Allen et al., 1989) and gap detection (Wightman et al., 1989) develop throughout childhood (for review see Sanes and Woolley, 2011) and contribute to gradual improvements in speech-in-noise perception with age. Improving young children's speech comprehension in degraded listening conditions may confer developmental advantages by elevating a child's learning potential.

While older children and adults with music training backgrounds demonstrate better speech perception in noise than untrained counterparts (Parbery-Clark et al., 2009b, 2011a; Strait et al., 2012; Zendel and Alain, 2011),

\* Corresponding author at: Auditory Neuroscience Laboratory, 2240 Campus Drive, Evanston, IL 60208, USA. Tel.: +1 847 491 2465; fax: +1 847 467 1464.

E-mail addresses: [dstrait@isr.umd.edu](mailto:dstrait@isr.umd.edu) (D.L. Strait), [apc@u.northwestern.edu](mailto:apc@u.northwestern.edu) (A. Parbery-Clark), [s-connell@northwestern.edu](mailto:s-connell@northwestern.edu) (S. O'Connell), [nkraus@northwestern.edu](mailto:nkraus@northwestern.edu) (N. Kraus).

<sup>1</sup> Now at: Neural Systems Laboratory, Institute for Systems Research, A. James Clark School of Engineering, University of Maryland, College Park, MD, USA.

**Table 1**

Description of musicians' training backgrounds. Each activity's administration style is specified as either "group" or "private."

	Music activity 1	Music activity 2	Music activity 3
1	Baby Steps (group)	Orff (group)	
2	Baby Steps (group)	Suzuki guitar (private)	
3	Orff (group)	Preschool music class (group)	
4	Music Together (group)	Drumming (group)	
5	Music Makers (group)	Musikgarten (group)	Preschool music camp (group)
6	Orff (group)	Voice lessons (private)	
7	Kindermusik (group)	Suzuki violin (private)	
8	Kindermusik (group)	Suzuki cello (private)	
9	Suzuki piano (private)	Suzuki cello (private)	
10	Suzuki violin (private)	Wiggleworms (group)	Mommy and Me (group)
11	Handbells (group)	Drumming (group)	
12	Wiggleworms (group)	Kindermusik (group)	Preschool music class (group)
13	Wiggleworms (group)	Music Together (group)	Preschool music class (group)
14	Suzuki piano (private)	Preschool music class (group)	
15	Suzuki piano (private)	Preschool music class (group)	
16	Orff (group)	Suzuki cello (private)	
17	Kindermusik (group)	Suzuki piano (private)	
18	Suzuki piano (private)		

concurrent with less of a detrimental impact of noise on auditory brainstem responses to speech (Parbery-Clark et al., 2009a, 2012b; Strait et al., 2012), we do not know if these advantages can be observed during early childhood. Although potentially impactful at any age, music training initiated during the first few years of life takes place alongside, and likely interacts with, extensive structural maturation within the auditory system. The auditory brainstem, for instance, undergoes significant postnatal myelination and synaptic refinement throughout the first two years of life (Moore and Linthicum, 2007; Moore et al., 1995). Developmental changes in speech-evoked auditory brainstem function that have been observed beyond the age of two may be attributed to ongoing cortical maturation, which extends into adolescence (Gleeson and Walsh, 2000; Moore and Guan, 2001; Moore and Linthicum, 2007), by means of descending cortico-collicular projections (Johnson et al., 2008).

Music training during the first few years of life has further capacity to shape fundamental characteristics of auditory function given the mammalian auditory system's amenability to learning. In fact, the modulation of auditory function according to experience with sound begins not long after the onset of hearing: newborn infants respond to their mothers' voices at lower levels than the voices of others, which is thought to reflect many weeks of experience listening to that voice in utero (DeCasper and Fifer, 1980). A child's interaction with musical sound during early childhood may promote the development of a more resilient auditory system by providing an acoustically-enriched auditory environment, with the potential to promote synaptic diversity (Bose et al., 2010), faster neural timing (Engineer et al., 2004; Jakkamsetti et al., 2012), frequency tuning (Jakkamsetti et al., 2012; Zhang et al., 2001) and to improve the responsiveness of neurons within auditory tracts (Alladi et al., 2005; Engineer et al., 2004; Percaccio et al., 2007).

Here, we aimed to determine the neural impact of music training initiated during the first four years of life by comparing auditory brainstem responses (ABRs) in musically-trained and untrained preschoolers who were

matched according to a variety of overarching factors (e.g., socioeconomic status, sex, IQ) to speech in quiet and noisy backgrounds. We assessed ABRs according to standard metrics known to relate to speech-in-noise perception in children, including the timing of discrete peaks in the response and the magnitude of the response to the stimulus' fundamental frequency and higher harmonics (Anderson et al., 2010a, 2010b; Strait et al., 2012). We further addressed training-related effects by assessing longitudinal data in a subset of musicians and nonmusicians. We hypothesized that music training during early childhood shapes fundamental characteristics of the brain's response to speech in noise and confers protective benefits against the deleterious effects of noise and that one year of continued music training strengthens these effects.

## 2. Materials and methods

### 2.1. Participants

All experimental procedures were approved by the Northwestern University Institutional Review Board. Thirty-two normal hearing children ( $\leq 20$  dB pure tone thresholds at octave frequencies from 250 to 8000 Hz) between the ages of 3 and 5 participated in this study. Participants and legal guardians provided informed assent and consent, respectively, and participants were monetarily compensated for their time. No parent reported a participant history of neurological or developmental abnormalities. Inclusionary criteria included normal wave V click-evoked ABR latencies and normal verbal IQ as measured by the *Peabody Picture Vocabulary Test* (Dunn and Dunn, 1997).

Subjects were categorized as musicians (Mus,  $N = 18$ ) or nonmusicians (NonMus,  $N = 14$ ). Musicians were currently undergoing private or group music training (e.g., Kindermusik, Music Together, Orff music classes; see Table 1) and had done so for a minimum of twelve consecutive months leading up to the test date (duration of music training  $M = 2.4$  years,  $SD = 1.42$ ; training initiation age

$M = 1.90$  years,  $SD = 1.48$ ). Although the style of training varied across participants (e.g., group or private lessons, training on one or many instruments, focus on tonal or percussive instruments, incorporation of song and dance), all musicians attended classes weekly and used at-home practice materials at least four days per week. Preschool nonmusicians had no music-training exposure during the year leading up to the test and  $\leq 6$  months accumulated over the course of their lives. Only three nonmusicians had some degree of music training, consisting of three ( $N = 2$ ) or six ( $N = 1$ ) months of group music classes. Mus and NonMus did not differ according to age, sex, verbal IQ, click-ABR wave V latency, extent of extracurricular activity (indexed by average hours per year since birth) or socioeconomic status as inferred from maternal education (all  $p > 0.2$ ; see Stevens et al. (2009) for discussion regarding the predictive value of maternal education for inferring a child's socioeconomic status). Nonmusicians' extracurricular activities included Gymboree, day camps, school- or community-center coordinated programs and group physical activities, among others.

Thirteen participants returned for testing one year  $\pm 1$  month following their first session and were re-administered the same protocol (Mus = 7, NonMus = 6). Again, Mus and NonMus did not differ according to age, sex, verbal IQ, or socioeconomic status as inferred from maternal education (all  $p > 0.1$ ). This opportunity was only given to participants whose music training statuses had not changed since first test (e.g., musicians whose training did not lapse over the course of the year). Year two data were compared to those collected at the first test session in order to assess the impact of an additional year of music training compared to an absence of music lessons.

## 2.2. Auditory brainstem response recordings

### 2.2.1. Stimulus and conditions

The evoking stimulus was a six-formant, 170 ms speech syllable /da/ 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 amidst 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.

### 2.2.2. Recording parameters

Auditory brainstem responses were recorded to the speech sound /da/ in NeuroScan Acquire 4.3 (Compumedics) at a 20 kHz sampling rate 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. 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 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 with either a parent or a tester and watched movies of their choice, with soundtracks playing at  $\leq 40$  dB. This method has proven enormously successful for minimizing myogenic activity and maintaining subject alertness (Skoe and Kraus, 2010).

## 2.3. Data processing and analysis

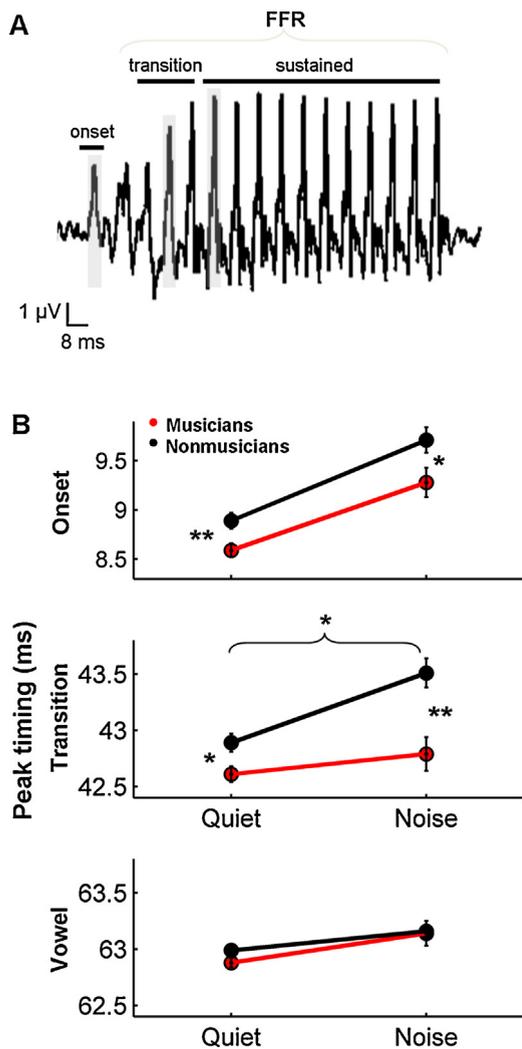
### 2.3.1. Data processing

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 falling within the phase-locking limits of auditory brainstem nuclei (Chandrasekaran and Kraus, 2010; Skoe and Kraus, 2010), epoched from  $-40$  to 190 ms referenced to the presentation of the stimulus (0 ms), and baseline corrected. Responses with amplitudes  $> \pm 35 \mu\text{V}$  were rejected as artifact. For each stimulus polarity 3000 artifact-free responses were averaged together and the resulting averages were added. All data processing was executed with scripts generated in MATLAB 7.5.0 (The Mathworks, Natick, MA).

The signal-to-noise ratio (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  $> 1.5$  in the quiet condition. Furthermore, musician and nonmusician groups were not distinct with regard to quiet response SNRs ( $F_{(1,31)} = 2.12$ ,  $p > 0.1$ ).

### 2.3.2. Auditory brainstem response timing and magnitudes

As in Strait et al. (2012), we gauged the effects of noise and musicianship on the neural response by identifying three response peaks corresponding to the onset of the neural response (9 ms), the onset of the formant transition of the stimulus (43 ms) and the onset of the sustained vowel region (63 ms) (Fig. 1A). Peaks yielded information on response timing (i.e., peak latencies) and magnitudes (i.e., peak amplitudes). 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



**Fig. 1.** Preschoolers with music training have faster auditory brainstem responses to speech. (A) Average auditory brainstem response to /da/ in quiet across all participants. Peaks corresponding to the stimulus onset, formant transition and sustained vowel are highlighted. (B) Onset and formant transition peak timing was earlier in musicians (red) than non-musicians (black) in both quiet and noise conditions. Musicians also demonstrated less of a timing delay from quiet to noise for the region of the response that corresponds to the formant transition. \* $p < 0.05$ , \*\* $p < 0.01$ .

participants' group characteristics, was sought. All peaks were identifiable in all participants.

### 2.3.3. Stimulus-to-response fidelity

As in Strait et al. (2012), we quantified the effect of noise on the fidelity of the neural response to the sustained vowel (i.e., the section of the response that best resembles the waveform of the evoking stimulus) by cross-correlating the stimulus and response waveforms. The degree of similarity was calculated by shifting the vowel section of the stimulus (60–170 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 sustained response. This time lag (7–12 ms) was chosen because it encompasses the stimulus transmission delay

(from the ER-3 transducer and ear insert ~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 statistical analyses.

### 2.3.4. Spectral encoding

To assess the neural encoding of the stimulus spectrum, we applied a fast Fourier transform to the sustained vowel portion of the responses (60–170 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 its harmonics  $H_2$ – $H_8$  (200–800 Hz).

### 2.3.5. Statistical analysis

Response characteristics in quiet and noise were compared using a repeated measures ANOVA with musicianship as fixed factor and condition (quiet/noise) as the within-subject factor. Independent samples  $t$ -tests were employed to better define the effects observed. Normality for all data was confirmed by the Shapiro–Wilk test for equality. Non-parametric Mann–Whitney  $U$  tests were employed to compare changes in neural response characteristics across the first and second testing sessions in light of the restricted sample size (13 musicians and nonmusicians). All statistical analyses were performed using SPSS (SPSS Inc., Chicago, IL) and reflect two-tailed values.

## 3. Results

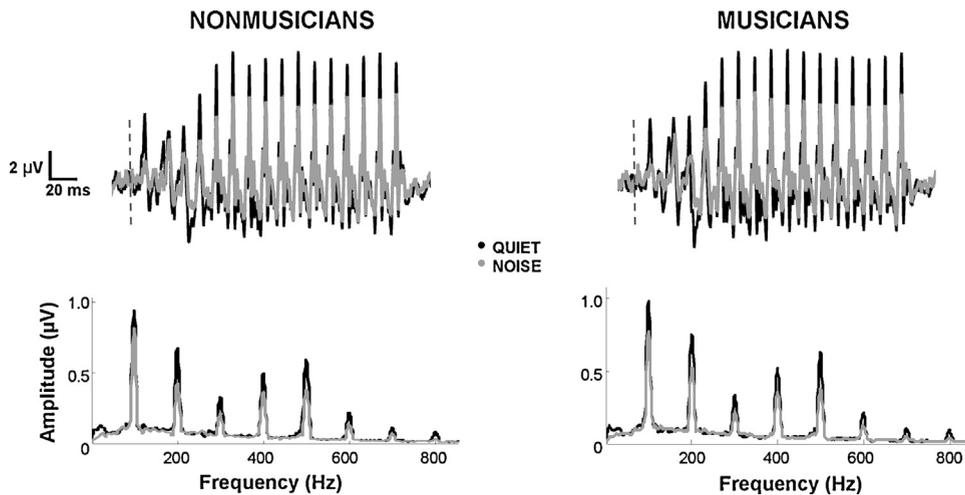
### 3.1. Summary of results

Musically trained children demonstrated faster neural responses to speech onsets and formant transitions in both quiet and noise conditions in addition to decreased quiet-to-noise timing delays and onset peak degradation. Musicians did not differ from nonmusicians on other neural response characteristics previously shown to differ between school-aged child and young adult musicians and nonmusicians (i.e., stimulus-to-response fidelity, spectral encoding of the higher harmonics). Test–retest data indicate that an additional year of continued music lessons further protects musicians' neural responses from the degrading effects of background noise.

### 3.2. Year one outcomes

#### 3.2.1. 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 (Fig. 1). A 2 (group)  $\times$  2 (condition)  $\times$  3 (peak) RMANOVA revealed significant main effects of group and noise, with responses in noise occurring later than responses in quiet ( $F_{(1,30)} = 74.37$ ,  $p < 0.0001$ ) and with musicians' responses occurring earlier than nonmusicians' ( $F_{(1,30)} = 9.61$ ,  $p = 0.004$ ). There was also a group  $\times$  condition  $\times$  peak interaction ( $F_{(3,28)} = 3.22$ ,  $p = 0.05$ ). Independent samples  $t$ -tests indicate that the



**Fig. 2.** Background noise degrades speech-evoked auditory brainstem responses, evident in temporal (upper panel) and spectral domains (lower panel). Preschoolers with and without music training demonstrated equivalent noise-induced degradation of auditory brainstem response stimulus-to-response correlations and spectral magnitudes.

onset and transition peaks occurred earlier in musicians than in nonmusicians in both quiet and noise (onset<sub>QUIET</sub>:  $t=2.40$ ,  $p=0.02$ ,  $d=0.87$ ; transition<sub>QUIET</sub>:  $t=2.29$ ,  $p=0.03$ ,  $d=0.83$ ; onset<sub>NOISE</sub>:  $t=2.33$ ,  $p=0.03$ ,  $d=0.85$ ; transition<sub>NOISE</sub>:  $t=3.07$ ,  $p=0.005$ ,  $d=1.12$ ). Musicians and nonmusicians did not differ in response timing to the sustained vowel (both  $t < 1.0$ ,  $p > 0.30$ ). With regard to quiet-to-noise timing shifts, we observed an interaction between noise and musician grouping ( $F_{(1,30)}=4.27$ ,  $p=0.04$ ) and independent samples  $t$ -tests revealed that musicians demonstrated less of a quiet-to-noise delay for the transition peak ( $t=2.44$ ,  $p=0.02$ ,  $d=0.85$ ). Group differences for the timing shift of the onset and vowel peaks did not approach significance ( $t \leq 1$ ,  $p > 0.3$ ).

### 3.2.2. Response magnitudes

We compared response magnitudes within conditions as well as the quiet-to-noise magnitude shifts for the onset, transition and sustained response peaks (occurring at  $\sim 9$ , 43 and 63 ms). A 2 (group)  $\times$  2 (condition)  $\times$  3 (peak) RMANOVA revealed a main effect of noise, with response magnitudes becoming smaller in noise ( $F_{(1,30)}=125.78$ ,  $p < 0.0001$ ), as well as a condition  $\times$  peak interaction ( $F_{(3,28)}=2.95$ ,  $p=0.05$ ). There was no main effect of group or a condition  $\times$  group  $\times$  peak interaction (see Section 3.3 for test-retest outcomes related to onset peak degradation with noise).

### 3.2.3. Stimulus-to-response fidelity and spectral encoding

A 2 (group)  $\times$  2 (condition) RMANOVA confirmed a significant main effect of noise on stimulus-to-response fidelity, with weaker correlations in noise relative to quiet ( $F_{(1,30)}=34.56$ ,  $p < 0.0001$ ). Similarly, a main effect of noise was observed on spectral encoding, with spectral magnitudes in noise decreased relative to quiet (2 (group)  $\times$  2 (condition)  $\times$  9 (spectral peak) RMANOVA  $F_{(1,30)}=62.98$ ,  $p < 0.0001$ ). Preschoolers with and without music training were not distinct with regard to stimulus-to-response

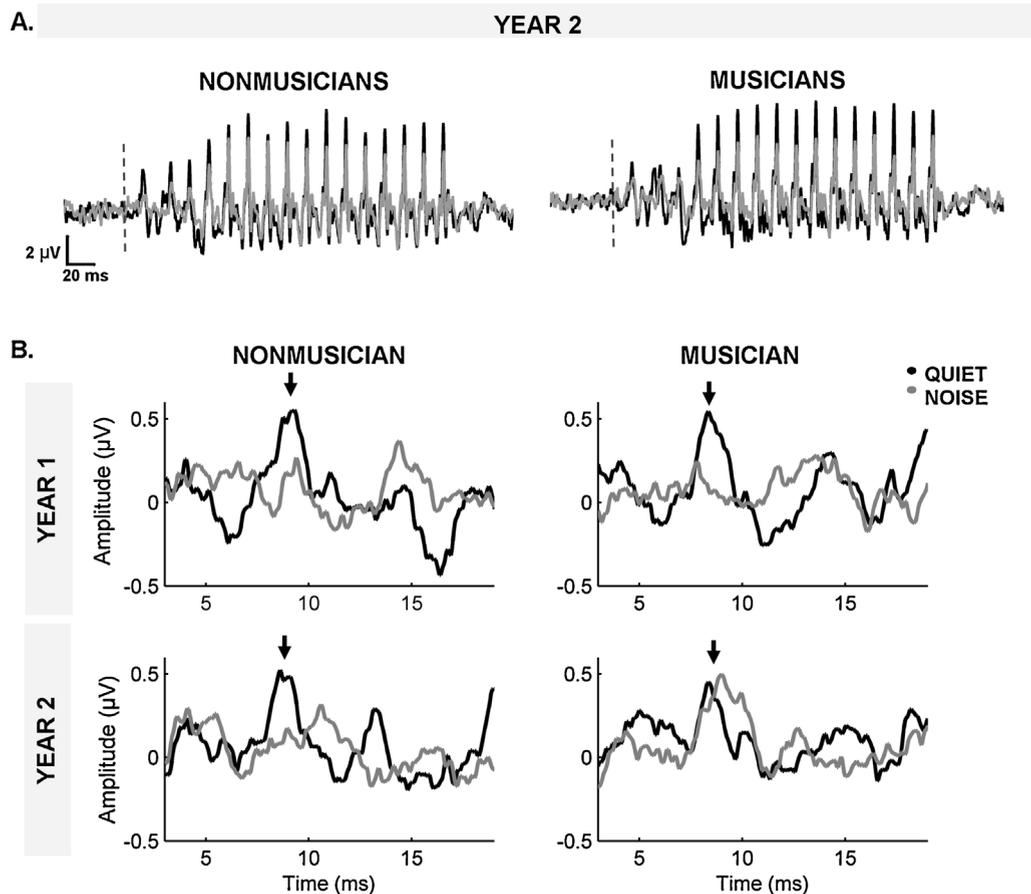
fidelity within either the quiet or the noise condition, nor in a comparison between the two. The same was observed for spectral magnitudes, including both the neural encoding of the fundamental frequency and higher harmonics (Fig. 2; all  $F < 1.0$ ,  $p > 0.4$ ).

### 3.3. Test-retest outcomes from years one and two

We assessed developmental changes over the course of one year in response magnitudes, timing, stimulus-to-response fidelity and spectral encoding in the subset of participants who returned for data collection  $\sim$ one year following their first test. Consistent with the test-retest stability of auditory brainstem responses in children (Hornickel et al., 2012), 2 (test session)  $\times$  2 (condition)  $\times$  3 (peak) and 2 (test session)  $\times$  2 (condition) RMANOVAs (assessing peak timing/magnitude and stimulus-to-response fidelity/spectral encoding, respectively) revealed no main effect of test session on any measure nor test session  $\times$  condition or peak or session  $\times$  condition  $\times$  peak interactions (all  $F < 1.2$ ,  $p > 0.3$ ). Differences in developmental changes between musicians and nonmusicians, however, indicated that preschoolers undergoing music training demonstrated more resilient onset responses to the presence of background noise with one further year of training than nonmusicians (i.e., smaller onset peak reductions from quiet to noise, Figs. 3B and 4; Year 1 onset peak degradation within these subjects only: Mann-Whitney  $U=0.72$ ,  $p=0.4$ ; Year 2 onset peak degradation Mann-Whitney  $U=3.00$ ,  $p=0.003$ ,  $d=3.14$ ; Year 1 to Year 2 change in onset peak degradation:  $U=2.00$ ,  $p=0.04$ ,  $d=0.92$ ).

## 4. Discussion

These data reveal that musicians' more resilient neural encoding of speech in noise arises early in life (i.e., as young as age three) and can be observed in children with as little as one year of training. Although musician preschoolers



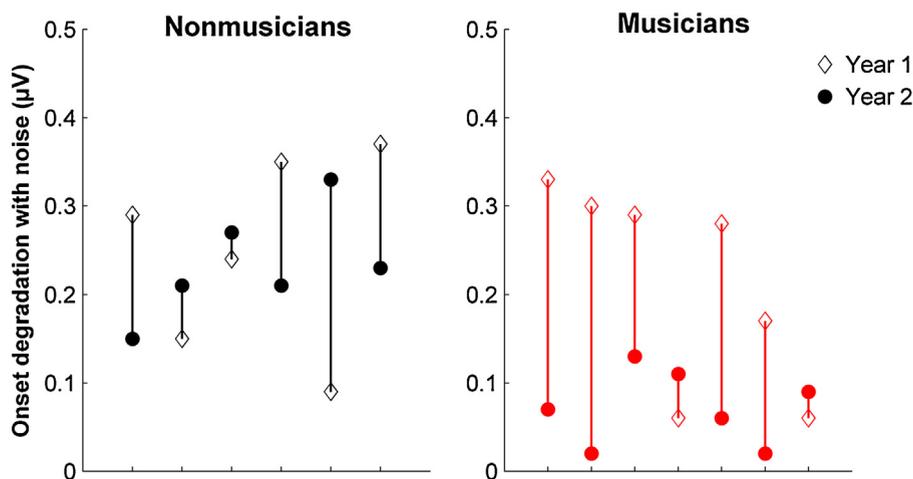
**Fig. 3.** (A) Auditory brainstem responses to the speech sound /da/ in quiet and noise for participants who returned for longitudinal data collection. (B) Auditory brainstem onset responses in two representative subjects. Musicians had more resilient onset responses to the presence of noise with an additional year of music training and development than nonmusicians.

have faster neural timing and less timing delays in noise for processing speech, however, they do not yet show the array of enhancements observed in older child (Strait et al., 2012) and adult musicians (Parbery-Clark et al., 2009a). Further training and development may be required for additional enhancements to emerge, especially with regard to the neural encoding of speech harmonics that is enhanced in older child and adult musicians. In fact, the neural encoding of speech harmonics is not only the last signature of musicianship to emerge during development but is also the first to fade in the aging process as hearing becomes compromised (Parbery-Clark et al., 2013). It is possible that refined neural timing mechanisms during development subsequently lead to more robust encoding of harmonics by increasing the precision with which neurons phase-lock to high-frequencies in speech. Their disappearance in aging musicians with compromised hearing, in contrast, may stem from the delayed timing and decreased neural response variability known to accompany increasing age (e.g., Anderson et al., 2012; Schmiedt et al., 1996; Turner et al., 2005).

It is important to note that, as in older children and adults, musician preschoolers' neural timing enhancements are exclusive to the region of the response that

corresponds to the neural encoding of the speech sound's onset and formant transition (i.e., the rapid transition from the consonant to the sustained vowel) and are not observed in response to the sustained vowel. Speech-in-noise perception consistently correlates with ABR onset and formant transition timing (Anderson et al., 2010a; Hornickel et al., 2009; Parbery-Clark et al., 2009a, 2012c; Strait et al., 2012), bringing to attention the many decades of psychophysical research establishing the importance (and difficulty) of stop-consonant perception during listening in noise (e.g., Miller and Nicely, 1955). While we have not yet determined preschool musicians' speech-in-noise abilities, based on these biological outcomes we would predict improved speech-in-noise perception in preschoolers undergoing music training.

In summary, our results indicate that musicians' strengthened neural speech-sound processing in noise is just emerging in this age group. Accordingly, longitudinal outcomes indicate an increased benefit of one additional year of continued music practice on neural responses to speech in noise, evidenced by less response magnitude degradation from quiet to noisy backgrounds. To our knowledge these data present the first evidence for a biological impact of infant/preschool music training, which



**Fig. 4.** Test–retest data across all subjects who returned for longitudinal data collection. Musicians demonstrated significant decreases in onset peak degradation by noise with an additional year of music training and development.

may support evidence for enhanced communication skills as a function of early music classes (Gerry et al., 2012). We interpret our findings according to the training-induced malleability of selective auditory brainstem response properties during a sensitive developmental period.

#### 4.1. Auditory experience during sensitive developmental periods may set the stage for later language and learning success

Early childhood constitutes a sensitive period during which auditory experience can have the most dramatic effect on neural structure and function (for review see Kral and Eggermont, 2007). Auditory input during the first five years of life contributes to a biological foundation for sound processing. Here, preschool musicians initiated consistent music training during their first four years of life; in fact, half of these children initiated training during the first year of life. Although these data provide the first evidence for the impact of music training on preschoolers' neural responses to speech, other forms of auditory experience during preschool years, such as language and reading exposure, have been associated with neural speech-sound sensitivity (Molfese et al., 2003). Musicians' early enhancements for speech processing may contribute to later auditory advantages given that neural responses to speech recorded during the first four years of life predict language performance later in childhood (Espy et al., 2004). This may account for why nonmusician adults with musically-active childhoods demonstrate auditory processing advantages that stem from their practice as children (Skoe and Kraus, 2012). In contrast, sound deprivation during early childhood, such as that caused by hearing loss, has lasting negative impacts on the auditory system. This is especially the case when sound deprivation occurs during the first 4–5 years of life (Schilder et al., 1994), after which interventions (e.g., cochlear implantation) have a comparably minimal impact on language development (Fryauf-Bertschy et al., 1997) (for review see Kral and Sharma, 2012; Sharma and Campbell, 2011).

In addition to shaping neural development, auditory enrichment during the first few years of life has the capacity to prime the auditory system for later learning success. This has been directly observed in animal models with auditory enrichment: rats exposed to music shortly after the onset of hearing, for example, demonstrate enhanced NMDA receptor (NMDAR) subunit NR2B protein expression in auditory cortex (Xu et al., 2009), which facilitates the experience-dependent synaptic plasticity that undergirds learning (for review see Bliss and Collingridge, 1993). Early exposure to auditory enrichment is crucial for these changes to occur: the identical enrichment strategy initiated following the onset of puberty confers no measurable effect on NMDA receptor proteins. Inverse relationships have been observed for post-natally sound-deprived rats, who have deficient NMDAR subunit expression both within auditory cortex (Bi et al., 2006) as well as in the auditory brainstem (Marianowski et al., 2000). Taken together, this suggests that early exposure to auditory-rich, interactive and engaging environments during initial development, such as that facilitated by music training, may set the stage for later auditory learning success such as musicians' enhanced learning of a second language (Shook et al., 2013), picking up on sound patterns (Schon and Francois, 2011), phonological awareness (Escalda et al., 2011) and reading acquisition (Douglas and Willatts, 1994; Moreno et al., 2009).

The cognitive emphasis of music training, such as on memory for and attention to sound, may provide neural benefits beyond those associated with acoustic enrichment. This emphasis could contribute to enhanced cognitive performance later in life, as seen in older child and adult musicians (Chan et al., 1998; George and Coch, 2011; Ho et al., 2003; Jakobson et al., 2008; Kraus et al., 2012; Pallesen et al., 2010; Parbery-Clark et al., 2009b, 2011a; Strait et al., 2010, 2012) and longitudinally in children who undergo music training (Moreno et al., 2011). Early cognitive stimulation may further promote the development of language and communication skills (Cates et al., 2012). Further efforts should assess the relationship of preschool

musicians' subcortical speech processing enhancements to cognitive performance, including speech-in-noise perception, memory, attention and language abilities.

While the data presented here indicate emerging biological advantages for processing speech in noise in musically trained preschoolers, how these advantages are reflected in everyday interactions with speech remains undefined. Future studies might assess speech and speech-in-noise perception in musically trained and untrained preschoolers, such as through the application of Litovsky's Children's Realistic Index of Speech Perception (CRISP), which can be administered to children as young as age three (Garadat and Litovsky, 2007).

#### 4.2. Longitudinal evidence for music training's impact on subcortical auditory function

Accumulating evidence substantiates subcortical auditory processing advantages in musicians compared to nonmusicians (Bidelman et al., 2009; Bidelman and Krishnan, 2010; reviewed in Kraus and Chandrasekaran, 2010; Musacchia et al., 2007; Parbery-Clark et al., 2009a, 2011b, 2012a, 2012c; Strait et al., 2009; Wong et al., 2007), observed as young as the school-age years (Strait et al., 2012). Effects of training have been indicated by consistent correlations between extent of music training and auditory brainstem response properties (Musacchia et al., 2007; Parbery-Clark et al., 2011b; Skoe and Kraus, 2012; Strait et al., 2009; Wong et al., 2007), including responses to speech in noise (Strait et al., 2012). Still, interpretations have been limited by between-groups designs that cannot disentangle experience-related from innate aspects of musicians' auditory processing enhancements. Although structural and functional effects of music training have been demonstrated longitudinally in the cortex, no studies have directly assessed the impact of music training on subcortical auditory function longitudinally.

Here, we present the first longitudinal evidence for music training's impact on auditory brainstem responses to speech: preschoolers undergoing music training diverge even further from nonmusicians according to the impact of noise on auditory brainstem responses after one year. Furthermore, group comparisons reveal that preschoolers with music training do not demonstrate all of the enhancements observed in older child and adult musicians (Parbery-Clark et al., 2009a; Strait et al., 2012); although they have faster response timing and less quiet-to-noise delays, no differences were observed between musicians' and nonmusicians' spectral magnitudes or stimulus-to-response fidelity. These aspects of musicians' speech-in-noise processing enhancements may not emerge without more extensive training.

Although longitudinal assessments of short-term auditory training programs have revealed that the speech-evoked auditory brainstem response is malleable with interactive auditory experience, including responses to speech in noise (for review see Kraus and Hornickel, 2012; Russo et al., 2005, 2010; Song et al., 2012), we are the first to indicate music training-related changes. In addition, we are the first to show training-related changes in brainstem onset response properties. An onset advantage has

previously been reported in musician adults, who experience less noise-induced onset peak degradation than nonmusicians (Parbery-Clark et al., 2009a, 2012b). The malleability of this obligatory response with music training may bear significant behavioral advantages, especially for speech perception (Parbery-Clark et al., 2009a, 2012b). Music also confers advantages beyond short-term training programs, including its inherently engaging, emotionally-rewarding characteristics. Music training provides access to a limitless variety of auditory input and training techniques that can be tailored to accommodate an individual's interests, strengths and weaknesses. Furthermore, early childhood music classes are often group-administered and consequently encourage the development of healthy social and communication skills (Gerry et al., 2012), in contrast to the computer delivery required for many short-term approaches that largely depends on individual administration. Music may provide a naturalistic mechanism for auditory training that is not only easy to implement but that can be initiated prior to a child's readiness to learn from more structured or computerized methods.

While these data indicate longitudinal outcomes of music training on the neural processing of speech in noise, we do not rule out the contribution of innate predispositions distinguishing children who undergo music training from those who do not. The longitudinal outcome presented here indicates that, even in the potential presence of innate predispositions, noise-induced auditory brainstem response degradation can be lessened with one year of continued music practice. Future studies that capture musicians prior to the onset of training are critical in order to adequately disentangle experience-related and innate factors.

#### Conflict of interest statement

The authors declare no competing financial interests.

#### Acknowledgments

This research was funded by the National Institutes of Health grant F31DC011457-01 to D.S., the National Science Foundation grant BCS-0921275 to N.K. and the Knowles Hearing Center. The authors would like to thank Karen Chan, Jessica Slater and Victor Abecassis for their assistance with data collection and Emily Hittner for her contributions to data reduction.

#### References

- Alladi, P.A., Roy, T., Singh, N., Wadhwa, S., 2005. Prenatal auditory enrichment with species-specific calls and sitar music modulates expression of Bcl-2 and Bax to alter programmed cell death in developing chick auditory nuclei. *International Journal of Developmental Neuroscience* 23, 363–373.
- Allen, P., Wightman, F., Kistler, D., Dolan, T., 1989. Frequency resolution in children. *Journal of Speech and Hearing Research* 32, 317–322.
- Anderson, S., Chandrasekaran, B., Skoe, E., Kraus, N., 2010a. Neural timing is linked to speech perception in noise. *Journal of Neuroscience* 30, 4922–4926.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., Kraus, N., 2012. Aging affects neural precision of speech encoding. *Journal of Neuroscience* 32, 14156–14164.

- Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., Kraus, N., 2010b. Brainstem correlates of speech-in-noise perception in children. *Hearing Research* 270, 151–157.
- Bi, C., Cui, Y., Mao, Y., Dong, S., Zhang, J., Sun, X., 2006. The effect of early auditory deprivation on the age-dependent expression pattern of NR2B mRNA in rat auditory cortex. *Brain Research* 1110, 30–38.
- 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, 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.
- Bliss, T.V., Collingridge, G.L., 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361, 31–39.
- Bose, M., Munoz-Llancao, P., Roychowdhury, S., Nichols, J.A., Jakkamsetti, V., Porter, B., Byrapureddy, R., Salgado, H., Kilgard, M.P., Aboitiz, F., Dagnino-Subiabre, A., Atzori, M., 2010. Effect of the environment on the dendritic morphology of the rat auditory cortex. *Synapse* 64, 97–110.
- Bradley, J.S., Sato, H., 2008. The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America* 123, 2078–2086.
- Cates, C.B., Dreyer, B.P., Berkule, S.B., White, L.J., Arevalo, J.A., Mendelsohn, A.L., 2012. Infant communication and subsequent language development in children from low-income families: the role of early cognitive stimulation. *Journal of Developmental and Behavioral Pediatrics* 33, 577–585.
- Chan, A.S., Ho, Y.C., Cheung, M.C., 1998. Music training improves verbal memory. *Nature* 396, 128.
- Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47, 236–246.
- DeCasper, A.J., Fifer, W.P., 1980. Of human bonding: newborns prefer their mothers' voices. *Science* 208, 1174–1176.
- Douglas, S., Willatts, P., 1994. The relationship between musical ability and literacy skills. *Journal of Research in Reading* 17, 99–107.
- Dunn, L.M., Dunn, L.M., 1997. *Peabody Picture Vocabulary Test*, third edition. Pearson Education, Inc., San Antonio, TX.
- 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, 651–653.
- Engineer, N.D., Percaccio, C.R., Pandya, P.K., Moucha, R., Rathbun, D.L., Kilgard, M.P., 2004. Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons. *Journal of Neurophysiology* 92, 73–82.
- Escalda, J., Lemos, S.M., Franca, C.C., 2011. Auditory processing and phonological awareness skills of five-year-old children with and without musical experience. *Jornal da Sociedade Brasileira de Fonoaudiologia* 23, 258–263.
- Espy, K.A., Molfese, D.L., Molfese, V.J., Modglin, A., 2004. Development of auditory event-related potentials in young children and relations to word-level reading abilities at age 8 years. *Annals of Dyslexia* 54, 9–38.
- 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, 3023–3029.
- Fryauf-Bertschy, H., Tyler, R.S., Kelsay, D.M., Gantz, B.J., Woodworth, G.G., 1997. Cochlear implant use by prelingually deafened children: the influences of age at implant and length of device use. *Journal of Speech, Language, and Hearing Research* 40, 183–199.
- Garadat, S.N., Litovsky, R.Y., 2007. Speech intelligibility in free field: spatial unmasking in preschool children. *Journal of the Acoustical Society of America* 121, 1047–1055.
- George, E.M., Coch, D., 2011. Music training and working memory: an ERP study. *Neuropsychologia* 49, 1083–1094.
- Gerry, D., Unrau, A., Trainor, L.J., 2012. Active music classes in infancy enhance musical, communicative and social development. *Developmental Science* 15, 398–407.
- Gleeson, J.G., Walsh, C.A., 2000. Neuronal migration disorders: from genetic diseases to developmental mechanisms. *Trends in Neurosciences* 23, 352–359.
- Hall 3rd, J.W., 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, 159–165.
- Ho, Y.C., Cheung, M.C., Chan, A.S., 2003. Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychologia* 17, 439–450.
- Hornickel, J., Knowles, E., Kraus, N., 2012. Test–retest consistency of speech-evoked auditory brainstem responses in typically-developing children. *Hearing Research* 284, 52–58.
- 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 United States of America* 106, 13022–13027.
- Jakkamsetti, V., Chang, K.Q., Kilgard, M.P., 2012. Reorganization in processing of spectral and temporal input in the rat posterior auditory field induced by environmental enrichment. *Journal of Neurophysiology* 107, 1457–1475.
- 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, 41–55.
- Johnson, K.L., Nicol, T., Zecker, S.G., Kraus, N., 2008. Developmental plasticity in the human auditory brainstem. *Journal of Neuroscience* 28, 4000–4007.
- Klatt, D., 1980. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67, 13–33.
- Kral, A., Eggermont, J.J., 2007. What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews* 56, 259–269.
- Kral, A., Sharma, A., 2012. Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences* 35, 111–122.
- Kraus, N., Chandrasekaran, B., 2010. Music training for the development of auditory skills. *Nature Reviews Neuroscience* 11, 599–605.
- Kraus, N., Hornickel, J., 2012. Meaningful engagement with sound for strengthening communication skills. In: Geffner, D., Ross-Swain, D. (Eds.), *Auditory Processing Disorders: Assessment, Management and Treatment*, second edition.
- Kraus, N., Strait, D.L., Parbery-Clark, A., 2012. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Annals of the New York Academy of Sciences* 1252, 100–107.
- Litovsky, R.Y., 1997. Developmental changes in the precedence effect: estimates of minimum audible angle. *Journal of the Acoustical Society of America* 102, 1739–1745.
- Marianowski, R., Liao, W.H., Van Den Abbeele, T., Fillit, P., Herman, P., Frachet, B., Huy, P.T., 2000. Expression of NMDA, AMPA and GABA(A) receptor subunit mRNAs in the rat auditory brainstem. I. Influence of early auditory deprivation. *Hearing Research* 150, 1–11.
- Miller, G.A., Nicely, P.E., 1955. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27, 338–352.
- Molfese, D.L., Molfese, V.J., Key, A.F., Kelly, S.D., 2003. Influence of environment on speech-sound discrimination: findings from a longitudinal study. *Developmental Neuropsychology* 24, 541–558.
- Moore, J.K., Guan, Y.L., 2001. Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology* 2, 297–311.
- Moore, J.K., Linthicum Jr., F.H., 2007. The human auditory system: a timeline of development. *International Journal of Audiology* 46, 460–478.
- Moore, J.K., Perazzo, L.M., Braun, A., 1995. Time course of axonal myelination in the human brainstem auditory pathway. *Hearing Research* 87, 21–31.
- 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, 1425–1433.
- 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, 712–723.
- Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15894–15898.
- Pallesen, K.J., Brattico, E., Bailey, C.J., Korvenoja, A., Koivisto, J., Gjedde, A., Carlson, S., 2010. Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE* 5, e11120.
- Papso, C.F., Blood, I.M., 1989. Word recognition skills of children and adults in background noise. *Ear and Hearing* 10, 235–236.
- Parbery-Clark, A., Anderson, S., Hittner, E., Kraus, N., 2012a. Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging* 33, 1483 e1481–1484.
- 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. *Frontiers in Aging Neuroscience* 4, 1–12.

- Parbery-Clark, A., Anderson, S., Kraus, N., 2013. Musicians change their tune: how hearing loss alters the neural code. *Hearing Research* 302, 121–131.
- Parbery-Clark, A., Skoe, E., Kraus, N., 2009a. Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience* 29, 14100–14107.
- Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009b. Musician enhancement for speech-in-noise. *Ear and Hearing* 30, 653–661.
- Parbery-Clark, A., Strait, D.L., Anderson, S., Hittner, E., Kraus, N., 2011a. Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS ONE* 6, e18082.
- Parbery-Clark, A., Strait, D.L., Kraus, N., 2011b. 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, D.L., Kraus, N., 2012c. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219, 111–119.
- Percaccio, C.R., Pruetter, A.L., Mistry, S.T., Chen, Y.H., Kilgard, M.P., 2007. Sensory experience determines enrichment-induced plasticity in rat auditory cortex. *Brain Research* 1174, 76–91.
- 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, 95–103.
- Sanes, D.H., Woolley, S.M., 2011. A behavioral framework to guide research on central auditory development and plasticity. *Neuron* 72, 912–929.
- Schilder, A.G., Snik, A.F., Straatman, H., van den Broek, P., 1994. The effect of otitis media with effusion at preschool age on some aspects of auditory perception at school age. *Ear and Hearing* 15, 224–231.
- Schmiedt, R.A., Mills, J.H., Boettcher, F.A., 1996. Age-related loss of activity of auditory-nerve fibers. *Journal of Neurophysiology* 76, 2799–2803.
- Schon, D., Francois, C., 2011. Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in Psychology* 2, 167.
- Sharma, A., Campbell, J., 2011. A sensitive period for cochlear implantation in deaf children. *Journal of Maternal-Fetal and Neonatal Medicine* 24 (Suppl. 1), 151–153.
- Shook, A., Marian, V., Bartolotti, J., Schroeder, S., 2013. Musical experience influences novel language learning. *American Journal of Psychology* 126, 95–104.
- Sininger, Y.S., Abdala, C., Cone-Wesson, B., 1997. Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response. *Hearing Research* 104, 27–38.
- Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. *Ear and Hearing* 31, 302–324.
- 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, 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, 1677–1688.
- Song, J.H., Skoe, E., Banai, K., Kraus, N., 2012. Training to improve hearing speech in noise: biological mechanisms. *Cerebral Cortex* 22, 1180–1190.
- 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, 634–646.
- 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, 661–668.
- Strait, D.L., Parbery-Clark, A., Hittner, E., Kraus, N., 2012. Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language* 123, 191–201.
- Turner, J.G., Hughes, L.F., Caspary, D.M., 2005. Affects of aging on receptive fields in rat primary auditory cortex layer V neurons. *Journal of Neurophysiology* 94, 2738–2747.
- Wightman, F., Allen, P., Dolan, T., Kistler, D., Jamieson, D., 1989. Temporal resolution in children. *Child Development* 60, 611–624.
- 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, 420–422.
- Xu, J., Yu, L., Cai, R., Zhang, J., Sun, X., 2009. Early auditory enrichment with music enhances auditory discrimination learning and alters NR2B protein expression in rat auditory cortex. *Behavioural Brain Research* 196, 49–54.
- Zendel, B.R., Alain, C., 2011. Musicians experience less age-related decline in central auditory processing. *Psychology and Aging* 27, 410–417.
- Zhang, L.I., Bao, S., Merzenich, M.M., 2001. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience* 4, 1123–1130.