

## Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians

A. Parbery-Clark<sup>a,b</sup>, D.L. Strait<sup>a,c</sup>, N. Kraus<sup>a,b,c,d,e,\*</sup>

<sup>a</sup> Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL, USA<sup>1</sup>

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

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

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

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

### ARTICLE INFO

#### Article history:

Received 22 February 2011

Received in revised form 11 July 2011

Accepted 8 August 2011

#### Keywords:

Auditory

Brainstem

Musical training

Speech in noise

### ABSTRACT

Musical training strengthens speech perception in the presence of background noise. Given that the ability to make use of speech sound regularities, such as pitch, underlies perceptual acuity in challenging listening environments, we asked whether musicians' enhanced speech-in-noise perception is facilitated by increased neural sensitivity to acoustic regularities. To this aim we examined subcortical encoding of the same speech syllable presented in predictable and variable conditions and speech-in-noise perception in 31 musicians and nonmusicians. We anticipated that musicians would demonstrate greater neural enhancement of speech presented in the predictable compared to the variable condition than nonmusicians. Accordingly, musicians demonstrated more robust neural encoding of the fundamental frequency (i.e., *pitch*) of speech presented in the predictable relative to the variable condition than nonmusicians. The degree of neural enhancement observed to predictable speech correlated with subjects' musical practice histories as well as with their speech-in-noise perceptual abilities. Taken together, our findings suggest that subcortical sensitivity to speech regularities is shaped by musical training and may contribute to musicians' enhanced speech-in-noise perception.

© 2011 Elsevier Ltd. All rights reserved.

Human communication rarely occurs in optimal listening environments; rather, we are often surrounded by background noise. Despite the frequent presence of noise, humans are remarkably adept at disentangling target sounds from a complex soundscape. A key mechanism thought to underlie accurate perception in noise is the auditory system's ability to extract regularities from an ongoing acoustic signal (Strait, Hornickel, & Kraus, in press; Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Winkler, Denham, & Nelken, 2009). This is accomplished through the neural fine-tuning of responses based on a sound's predictability—a phenomenon that is observed in humans, non-human primates and rodents, underscoring its importance for auditory function throughout the animal kingdom (Baldeweg, 2006; Bendixen, Roeber, & Schröger, 2007; Bendixen, Schroger, & Winkler, 2009; Chandrasekaran et al., 2009; Dean, Harper, & McAlpine, 2005; Dean, Robinson, Harper, & McAlpine, 2008; Malmierca, Cristaudo, Perez-Gonzalez, & Covey, 2009; Perez-Gonzalez, Malmierca, &

Covey, 2005; Pressnitzer, Sayles, Micheyl, & Winter, 2008; Skoe & Kraus, 2010b; Wen, Wang, Dean, & Delgutte, 2009; Winkler et al., 2009). Neural adaptation to regularly occurring sounds does not require overt attention (Muller, Metha, Krauskopf, & Lennie, 1999; Pressnitzer et al., 2008; Webster, Kaping, Mizokami, & Duhamel, 2004), reinforcing the idea that this sensitivity to acoustic regularities is a basic tenet of auditory processing. Consistent with the notion that regularity detection is a basic function of the auditory system, failure to take advantage of acoustic regularities has been linked to poor music aptitude (Strait, Hornickel et al., in press) and language impairments such as developmental dyslexia (Ahissar, Lubin, Putter-Katz, & Banai, 2006; Chandrasekaran et al., 2009; Kujala et al., 2000; Kujala & Naatanen, 2001; Schulte-Korne, Deimel, Bartling, & Renschmidt, 1999) and specific language impairment (Evans, Saffran, & Robe-Torres, 2009).

Pitch is a particularly important form of regularity that the auditory system employs to promote auditory object formation and speaker identification (Baumann & Belin, 2010; Clarke & Becker, 1969; Kreiman, Gerratt, Precoda, & Berke, 1992), two key elements required for extracting speech from a noisy environment (Oxenham, 2008; Shinn-Cunningham & Best, 2008). Pitch is the perceptual correlate of a sound's fundamental frequency ( $F_0$ ), the slowest repeating periodic element of that sound. Increasing

\* Corresponding author at: Northwestern University, 2240 Campus Drive, Evanston IL, 60208, USA. Tel.: +1 847 491 2459; fax: +1 847 491 2523.

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

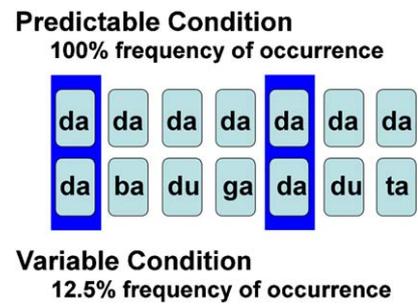
<sup>1</sup> <http://www.brainvolts.northwestern.edu>.

the pitch difference between two auditory streams improves a listener's ability to perceive them as two distinct auditory objects, resulting in better perception of their acoustic content (Assmann & Summerfield, 1987; Bird & Darwin, 1998; Brokx & Nooteboom, 1982; Culling & Darwin, 1993; Drullman & Bronkhorst, 2004). Indeed, these results highlight the important role that voice pitch plays for speech-in-noise perception. Recently, the relationship between  $F_0$  and speech-in-noise perception was extended to neurobiological processes whereby the robustness of an individual's subcortical encoding of the  $F_0$  correlates with speech-in-noise perceptual ability (Anderson, Skoe, Chandrasekaran, Zecker, & Kraus, 2010; Song, Skoe, Banai, & Kraus, 2010).

Music, like language, is a highly structured system that relies on the grouping of a finite number of sounds (notes) according to specific rules. Given that musical skill requires the rapid online detection of sound regularities, it is not surprising that musicians demonstrate advantages for detecting pitch patterns (Brattico, Näätänen, & Tervaniemi, 2001; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005). Musicians are also more adept at using rhythmic and tonal regularities to extract rhythmic hierarchies (Drake, Jones, & Baruch, 2000), to detect incongruous melodic endings (Besson & Faita, 1995; Besson, Faita, & Requin, 1994), and to structure temporal (Tervaniemi, Ilvonen, Karma, Alho, & Näätänen, 1997) and tonal sequences (van Zuijen et al., 2004). Evidence suggests that musicians' sensitivity to acoustic regularities extends to the domain of language, in which musicians are more sensitive to deviations in speech patterns than nonmusicians (Jentschke & Koelsch, 2009; Magne, Schön, & Besson, 2006; Nikjeh, Lister, & Frisch, 2009; Schon, Magne, & Besson, 2004; Tervaniemi et al., 2009). To date, this work has only focused on perceptual abilities and cortical processes. It remains to be determined whether the musician advantage for detecting sound patterns and acoustic regularities extends to other aspects of neural processing in the auditory pathway—namely, the auditory brainstem.

The auditory brainstem response (ABR) provides a noninvasive method to study subcortical sensitivity to acoustic regularities. While the ABR represents acoustic features of the evoking stimulus with considerable fidelity (Galbraith, Arbagey, Branski, Comerci, & Rector, 1995; Skoe & Kraus, 2010a), it is also modulated by stimulus predictability and the behavioural relevance of the evoking sound. Enhanced subcortical responses to meaningful acoustic features have been noted to occur on multiple time scales, such as over the course of a 90-min recording session (Skoe & Kraus, 2010b), with short-term auditory training (Russo, Nicol, Zecker, Hayes, & Kraus, 2005; Song, Skoe, Wong, & Kraus, 2008), and as a function of life-long musical or linguistic experience (Bidelman, Gandour, & Krishnan, 2011; Kraus & Chandrasekaran, 2010; Krishnan & Gandour, 2009; Krishnan, Xu, Gandour, & Cariani, 2005; Lee, Skoe, Kraus, & Ashley, 2009; Parbery-Clark, Skoe, & Kraus, 2009; Strait, Kraus, Skoe, & Ashley, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007). As such, the auditory brainstem offers a unique window through which we can explore the relative impacts of stimulus predictability and musical expertise on auditory processing and their influences on speech-in-noise perception.

Hearing in noise depends on the auditory system's ability to segregate a target signal from competing input. The pitch cues of a person's voice provide not only a means to distinguish between the target and the noise, but they also act as a unifying percept that enables elements with the same pitch to be grouped into a single acoustic stream. In fact, the ability to hear in noise relates with the magnitude of pitch, or  $F_0$ , encoding (Anderson et al., 2010; Song et al., 2010). Given this, we were surprised to find that musicians, who demonstrate a heightened ability to hear in noise, do not demonstrate enhanced subcortical  $F_0$  encoding for speech in noise compared to nonmusicians (Parbery-Clark, Skoe, &



**Fig. 1.** To investigate subcortical sensitivity to stimulus regularities, the speech sound /da/ was presented in predictable (top) and variable (bottom) conditions. In the predictable condition, /da/ was presented 100% of the time. In the variable condition, /da/ was randomly presented at a frequency of 12.5%, in the context of seven other speech sounds. Auditory brainstem responses to /da/ were event-matched between the two conditions (represented by highlighted bars) so as to avoid the potential confound of presentation order.

Kraus, 2009). While musicians' enhanced speech-in-noise perception (Parbery-Clark, Skoe, & Kraus, 2009; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011) suggests a potential advantage for locking on to the acoustic regularities of a target speaker's voice, mechanisms other than general  $F_0$  enhancement must underlie their heightened perceptual performance. Here, we ask whether musicians' speech-in-noise perception is founded on a greater neural sensitivity to acoustic regularities within an ongoing stream of speech sounds.

## 1. Methods

### 1.1. Participants

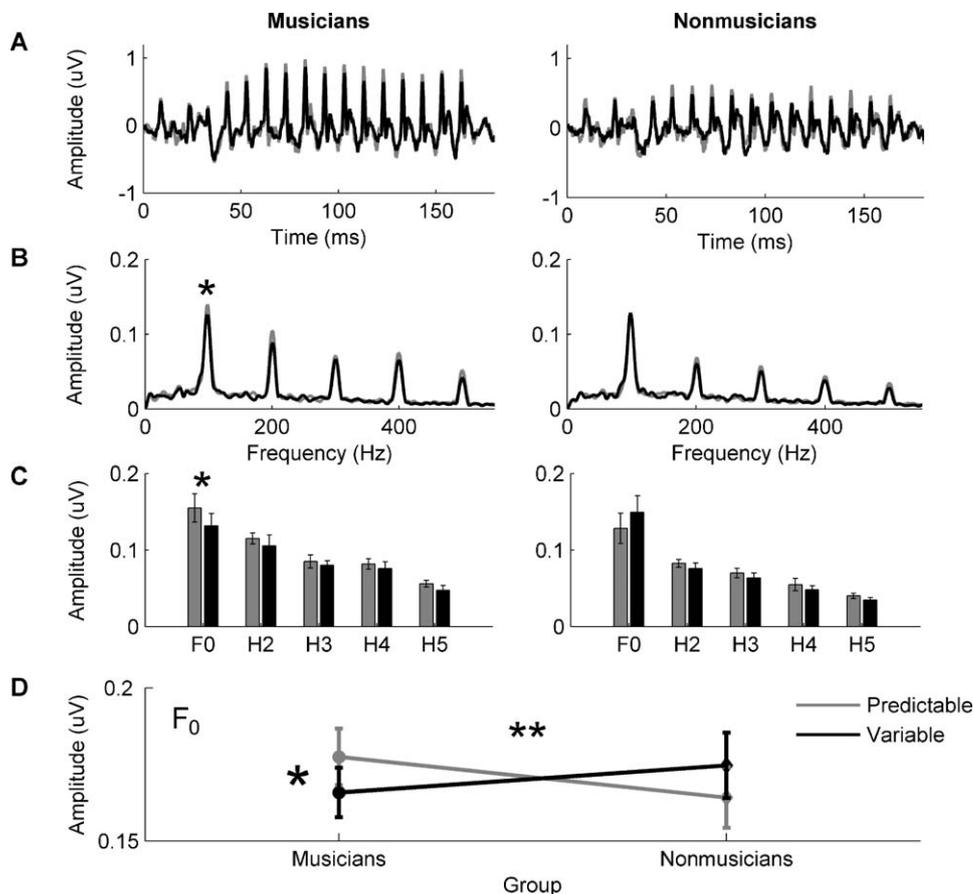
Thirty-one young adults ages 18–30 (mean = 22.4, s.d. = 3.4) participated in this study. All participants were native English speakers, had normal hearing thresholds (<20 dB HL at octave frequencies 0.125–8 kHz) and demonstrated normal IQ as measured by the Test of Nonverbal Intelligence (TONI) (Brown, Sherbenou, & Johnsen, 1982). No participants reported histories of learning or neurological disorders. All participants gave their written informed consent according to principles set forth by Northwestern University's Institutional Review Board. Sixteen participants were classified as musicians (11 females), all of whom had begun musical training at or before age seven (mean = 5.1, s.d. = 1.2) and had consistently played a musical instrument throughout their lives up until the time of study participation (mean = 16.4, s.d. = 3.4). Fifteen subjects were classified as nonmusicians (11 females) and reported fewer than three years of musical training at any point in their lives. The two groups did not differ according to IQ ( $F(1,29) = 2.25$ ,  $p = 0.144$ ), age ( $F(1,29) = 0.643$ ,  $p = 0.429$ ), or hearing thresholds ( $F(1,29) = 1.059$ ,  $p = 0.312$ ).

### 1.2. Stimuli

The speech stimulus was a 170 ms six-formant speech syllable (/da/) synthesized using a Klatt-based synthesizer at a 20 kHz sampling rate. This syllable has a steady fundamental frequency ( $F_0 = 100$  Hz) except for an initial 5 ms stop burst. During the first 50 ms (transition between the stop burst and the vowel), the first, second and third formants change over time ( $F_1$ , 400–720 Hz;  $F_2$ , 1700–1240 Hz;  $F_3$ , 2580–2500 Hz) but stabilize for the subsequent vowel, which is 120 ms in duration. The fourth, fifth and sixth formants are constant throughout ( $F_4$ , 3300 Hz;  $F_5$ , 3750 Hz;  $F_6$ , 4900 Hz). In the predictable condition, the /da/ was presented 100% of the time. In the variable condition, the same /da/ was randomly presented within the context of seven other speech sounds at a probability of 12.5% (Fig. 1). The seven other speech sounds varied by formant structure (/ba/, /ga/, /du/), voice-onset time (/ta/),  $F_0$  (/da/ with a dipping contour, /da/ with an  $F_0$  of 250 Hz), and duration (163 ms /da/) (see Chandrasekaran et al., 2009 for further descriptions of the stimuli and the experimental paradigm).

### 1.3. Auditory brainstem measures

Brainstem responses were differentially recorded at a 20 kHz sampling rate using Ag–AgCl electrodes in a vertical montage (Cz active, forehead ground and linked-earlobe references) using Scan 4.3 (Compumedics, Charlotte, NC) under the two different conditions (i.e., predictable and variable). In both conditions, a total of 6100 individual responses were collected. The speech stimuli were presented in alternating polarities, a common technique used in subcortical recordings to minimize the contribution of stimulus artifact and cochlear microphonic (Aiken & Picton, 2008; Gorga, Abbas, & Worthington, 1985). The speech stimuli were



**Fig. 2.** Subcortical enhancement of predictable speech in musicians. (A) Group average responses for musicians (left) and nonmusicians (right) in predictable (grey) and variable (black) stimulus conditions. (B) Grand average spectra and (C) corresponding frequency amplitude bar graphs for the predictable and variable conditions for the fundamental frequency ( $F_0$ ) and its harmonics ( $H_2$ – $H_5$ ). (D) Musicians, but not nonmusicians, demonstrate enhanced subcortical representation of the fundamental frequency ( $F_0$ ) in the predictable condition relative to the variable condition. \* $p < 0.05$ ; \*\* $p < 0.01$ ; error bars represent 1 standard error.

presented binaurally at 80 dB SPL with an inter-stimulus interval of 83 ms through insert ear phones (ER-3; Etymotic Research, Elk Grove Village, IL) at a rate of 3.95 Hz. Contact impedance was  $2\text{ k}\Omega$  or less across all electrodes. Each condition lasted  $\sim 26 \pm 2$  min. To facilitate a wakeful yet relaxed state for the recording session, participants watched a silent, captioned movie of their choice. Artifact rejection was monitored during the recording session and maintained at  $< 10\%$  for all participants.

Responses were band-pass filtered offline from 70 to 2000 Hz (12 dB roll-off, zero phase-shift) using NeuroScan Edit 4.3 to maximize the contribution of the auditory brainstem nuclei in the recording and to reduce the inclusion of low-frequency cortical activity (Akhoun et al., 2008; Chandrasekaran & Kraus, 2010). Responses were epoched from  $-40$  to 213 ms (stimulus onset occurring at 0 ms) and events with amplitudes beyond  $\pm 35 \mu\text{V}$  were rejected as artifact. To compare the neural response to /da/ across the two conditions, responses were trial matched (Fig. 1). This was achieved by ranking the presentation of the /da/ in the variable condition to its presentation number to create a presentation event template from which the responses to the /da/ at the same location in the predictable condition could be selected. The individual trials selected from each condition were then summed to create trial-matched averages. While this method results in a large number of trials being disregarded in the predictable condition, trial-matching between the two conditions facilitates the comparison of neural responses to the same stimulus recorded in two conditions without the potential confound of presentation order or differing numbers of trials. The final response average for each condition consisted of 700 artifact-free neural responses to /da/.

#### 1.4. Frequency analyses

The spectral energy of the neural responses was analyzed using MATLAB 7.5.0 (The Mathworks, Inc., Natick, MA) by computing fast Fourier transforms. Spectral amplitudes were calculated for the portion of the response corresponding to the vowel (60–170 ms) using 40 Hz-wide bins centered at the  $F_0$  (100 Hz) and its second through fifth harmonics (200–500 Hz). The differences between the spectral amplitudes of the  $F_0$  and its harmonics in the two conditions were then calculated by subtracting the spectral amplitudes of the frequencies of interest in the variable condition from the spectral amplitudes of these same frequencies in the predictable condition.

#### 1.5. Speech perception in noise

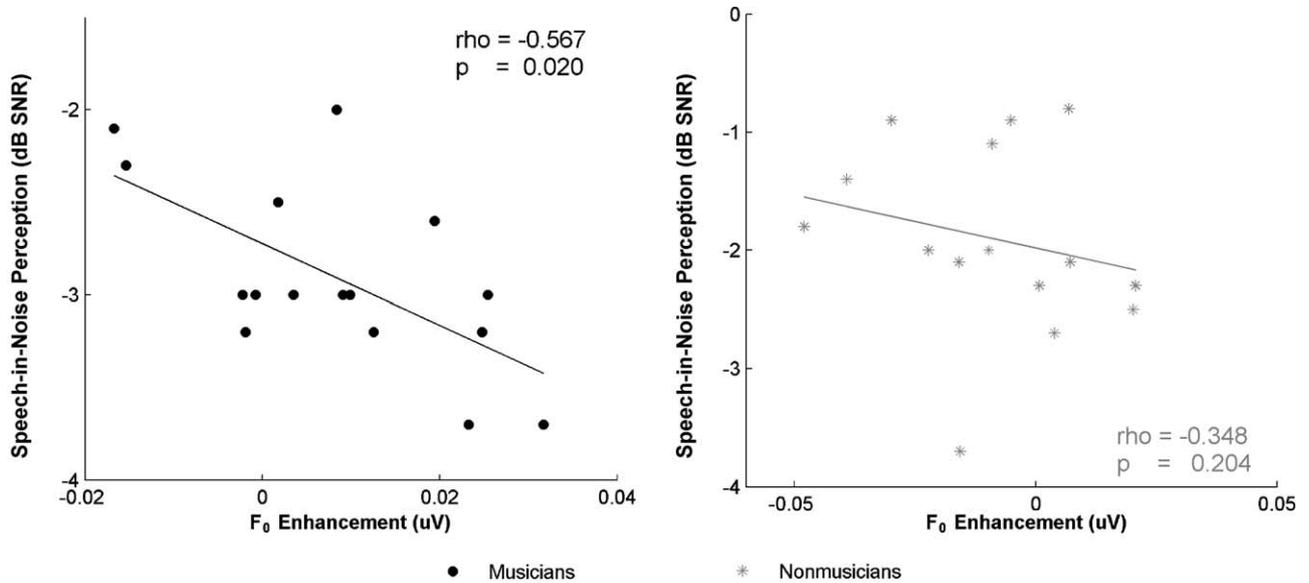
Speech perception in noise was measured using the Hearing in Noise Test (HINT, Bio-logic Systems Corp; Mundelein, IL) (Nilsson, Soli, & Sullivan, 1994). During this test participants repeat short, semantically and syntactically simple sentences (e.g., *she stood near the window*) taken from the Bamford–Kowal–Bench corpus of male-produced sentences (Bench, Kowal, & Bamford, 1979). These sentences are presented in free-field within speech-shaped background noise from a loud speaker located one meter at a  $0^\circ$  azimuth from the participant. The noise presentation level is fixed at 65 dB SPL and the program automatically adjusts perceptual difficulty by increasing or decreasing the intensity level of the target sentences until the signal-to-noise ratio (SNR) threshold is determined. This SNR threshold is defined as the dB level difference between the speech and the noise at which 50% of sentences are correctly repeated. A lower SNR indicates better performance.

#### 1.6. Statistical analyses

Group differences in the neural encoding of the  $F_0$  and its second through fifth harmonics ( $H_2$ – $H_5$ ) in the predictable and variable conditions were explored with a repeated measure ANOVA. To quantify relationships among stimulus predictability, speech-in-noise perception and musicians' musical practice histories, Pearson and Spearman correlations were employed as appropriate (SPSS, Chicago, IL). All results reflect two-tailed significance values.

## 2. Results

Consistent with previous results, musicians demonstrated better perception of speech in noise than nonmusicians ( $F(1,29) = 18.04$ ,  $p < 0.005$ ) (Parbery-Clark, Skoe, Lam et al., 2009). Musicians also showed greater subcortical enhancement of the  $F_0$  presented in the predictable compared to the variable condition than nonmusicians (Fig. 2). A RMANOVA (condition  $\times$  group)



**Fig. 3.** The difference between the  $F_0$  representation in the predictable compared to the variable condition ( $F_{0\text{predictable}} - F_{0\text{variable}}$ ) correlates with speech-in-noise perception among musicians, in which greater subcortical sensitivity to speech regularities relates to better speech-in-noise perception ( $\rho = -0.567$ ,  $p = 0.02$ ). This same relationship is not present in the nonmusician group ( $\rho = -0.348$ ,  $p = 0.204$ ).

demonstrated a significant interaction between stimulus condition and musician/nonmusician groups for the  $F_0$  ( $F(1,29) = 7.288$ ,  $p = 0.01$ ) but not for its harmonics ( $H_2$ – $H_5$ , all  $p > 0.5$ ). We did not observe a main effect of musical training on  $F_0$  encoding in that musicians did not demonstrate enhanced  $F_0$  encoding relative to nonmusicians within either the predictable or the variable conditions (predictable:  $F(1,30) = 0.659$ ,  $p = 0.423$ ; variable:  $F(1,30) = 0.375$ ,  $p = 0.545$ ). Post hoc, within-group, paired samples  $t$ -tests, however, revealed that musicians had a greater representation of the  $F_0$  in the predictable relative to the variable condition ( $t(15) = -2.336$ ,  $p = 0.034$ ) whereas nonmusicians did not ( $t(14) = 1.717$ ,  $p = 0.11$ ) (Fig. 2).

#### 2.1. Subcortical enhancement of predictable speech relates with speech-in-noise perception

Across all subjects, the degree of enhancement in auditory brainstem responses to the  $F_0$  between the two conditions correlated with speech-in-noise (SIN) perception, with better SIN perception relating to a greater  $F_0$  enhancement in the predictable compared

to the variable condition ( $r = -0.533$ ,  $p = 0.002$ ). Within-group analyses revealed that this relationship was mainly driven by the musicians ( $\rho = -0.567$ ,  $p = 0.02$ ) and was not significant among nonmusicians only ( $\rho = -0.348$ ,  $p = 0.204$ ) (Fig. 3). No other measure of spectral magnitude (i.e.,  $H_2$ – $H_5$ ) between the predictable and variable conditions related with SIN perception across all subjects or within each group separately ( $p > 0.2$ ).

#### 2.2. Subcortical enhancement of predictable speech relates with music experience

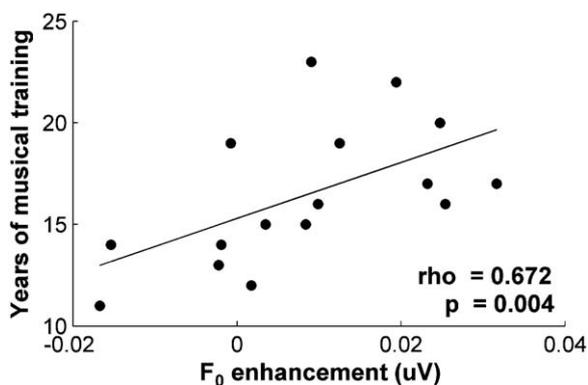
We observed a positive correlation between years of musical experience and the extent of  $F_0$  enhancement within musicians, with more years of practice relating to a greater enhancement of the  $F_0$  in the predictable relative to the variable condition (Fig. 4;  $\rho = 0.672$ ,  $p = 0.004$ ).

### 3. Discussion

Here, we demonstrate the facilitation of neural encoding according to regularities in an ongoing speech stream. In musicians, the representation of voice pitch in the auditory brainstem response is enhanced to predictably occurring speech relative to a variable, or unpredictable, context. This result highlights the potential role that musical training plays in improving the detection of acoustic regularities. Furthermore, the extent of subcortical enhancement to predictably occurring speech in musicians relates to the ability to hear speech in background noise. These findings demonstrate a neural advantage for detecting regularities within an ongoing speech stream in musicians, underscoring their enhanced ability to track the voice of a target speaker in background noise.

#### 3.1. Musical experience tunes neural sensitivity to acoustic regularities

Music, like language, is highly patterned, with both music and language demonstrating acoustic organization that unfolds over time. Within a musical context, repetition and regularity play fundamental roles, contributing to the perception of rhythm, meter (Hannon, Snyder, Eerola, & Krumhansl, 2004; Large & Jones, 1999)



**Fig. 4.** The difference between the  $F_0$  representations in the predictable compared to the variable condition ( $F_{0\text{predictable}} - F_{0\text{variable}}$ ) correlates with years of musical practice for the musician group. The more years musicians consistently practiced their instrument, the greater the neural sensitivity to speech sound regularities ( $\rho = 0.672$ ,  $p = 0.004$ ).

and tonality (Krumhansl, 1985, 1990). Given the common use of sound patterns in music and that musicians are trained to detect them, it is not surprising that musicians are more adept at distinguishing pitch (Brattico et al., 2001; van Zuijlen et al., 2004), rhythmic (van Zuijlen et al., 2005) and melodic (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005) patterns. In addition, musicians are more sensitive to harmonic relationships between chords, with musicians demonstrating greater effects of harmonic expectancy than nonmusicians (Koelsch, Jentschke, Sammler, & Mietchen, 2007).

In this study, we show that musicians' sensitivity to sound patterns transfers to the speech domain, with musicians demonstrating a greater representation of pitch cues in a predictable than a variable context relative to nonmusicians. Furthermore, we provide evidence that this skill is differentiated within the musician group in that musicians with more years of musical training show greater sensitivity to stimulus regularities. This correlation between the degree of subcortical sensitivity to stimulus regularities and years of musical practice provides compelling evidence that our results may be induced, at least in part, by training rather than being driven solely by genetic factors. Additional support for this premise comes from Moreno et al. (2009), who conducted a longitudinal study with random assignment of nonmusician children to music or painting training. Results showed that those children who received musical training demonstrated enhanced pitch processing for both music and speech, again highlighting the role of musical training, rather than genetic predispositions alone, in engendering musicians' enhanced processing of pitch. As such, our results contribute to a growing literature demonstrating a link between musicians' musical practice histories (Margulis, Mlsna, Uppunda, Parrish, & Wong, 2009; Pantev, Roberts, Schulz, Engelen, & Ross, 2001; Strait, Chan, Ashley, & Kraus, *in press*) and the number of years of musical practice or age of onset of musical training with the extent of neural enhancement observed (Gaser & Schlaug, 2003; Hutchinson, Lee, Gaab, & Schlaug, 2003; Lee et al., 2009; Musacchia, Sams, Skoe, & Kraus, 2007; Strait & Kraus, 2011; Strait et al., 2009; Wong et al., 2007).

### 3.2. Neural enhancement to acoustic regularities: the role of the descending auditory system

The brain is an active, adaptive system that generates predictions about the environment (Engel, Fries, & Singer, 2001; Friston, 2005; Winkler et al., 2009). Neural representation of the external world is not context-invariant; rather, experience (short- and long-term) and immediate environmental conditions modulate neural function. Out of a constant barrage of sensory information, cortical centers are thought to continually predict which features are behaviourally relevant and fine-tune the neural encoding of these features accordingly, thus facilitating their perception (Ahissar & Hochstein, 2004; Ahissar, Nahum, Nelken, & Hochstein, 2009; Bajo, Nodal, Moore, & King, 2010; Engel et al., 2001; Fritz, Elhilali, & Shamma, 2005; Tzounopoulos & Kraus, 2009). One way this may be accomplished is through the generation of predictions in cortical centers that are, in turn, used to bias neural encoding occurring earlier in the auditory processing stream. This process improves the representation of salient or predictable acoustic features. The modification of subcortical response properties is thought to occur via a series of reciprocal feedback loops originating in the cortex and terminating at subcortical nuclei (Suga & Ma, 2003). The descending branches of the auditory system modify subcortical responses by sharpening neural tuning and increasing response magnitudes for frequently occurring signals (Gao & Suga, 1998; Yan & Suga, 1998) or the most behaviourally relevant auditory information while suppressing irrelevant input (Gao & Suga, 1998; Luo, Wang, Kashani, & Yan, 2008).

Continually generated predictions also serve another important role: they enable the formation of templates that can be used to facilitate interpretation of incoming information. When the incoming stimulus matches the internally generated pattern, sensory processing can be accomplished more rapidly (i.e., because it was expected) (Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005). If, on the other hand, the incoming stimulus does not match the internally generated pattern, it is flagged as a violation of the predicted sound sequence. When this occurs, the auditory system must determine whether to modulate the current expectation or to treat the violation as a unique event (Näätänen & Winkler, 1999; Winkler, Karmos, & Näätänen, 1996). Our results indicate that musicians benefit from regularity to a greater extent than nonmusicians. This result highlights the potential presence of strengthened top-down control mechanisms in musicians, as has been previously proposed (Kraus & Chandrasekaran, 2010; Strait, Chan et al., *in press*; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Tervaniemi et al., 2009), that enable musicians to tailor subcortical structures for the maximal extraction and neural encoding of important acoustic information. An alternate but not mutually exclusive mechanism is that musical training may bolster stimulus-driven, bottom-up statistical learning. Increased modulation of subcortical processes according to statistical properties of incoming sounds can stem from top-down and/or bottom-up mechanisms and would result in an enhancement of regularities within a speech stream. While both of these mechanisms may be at play, the present data cannot disambiguate their relative contributions.

### 3.3. Subcortical sensitivity to acoustic regularities improves SIN perception

A key component of SIN perception is the ability to distinguish a target voice from background noise. The pitch of a target voice is a salient cue that contributes to the extraction of a voice from its surrounding acoustic environment (Assmann & Summerfield, 1987; Bird & Darwin, 1998; Brokx & Nootboom, 1982; Culling & Darwin, 1993; Drullman & Bronkhorst, 2004). Because the pitch of a person's voice is relatively constant and changes smoothly over time (Darwin, 2005), pitch cues are typically easy to follow. Pitch can also provide a means to group sequential sounds into an acoustic stream, contributing to object formation that facilitates speech perception in noise. Still, how does this translate into a neural code? One idea is that the auditory system enhances its representation of the target stimulus' spectral cues to positively contribute to object formation (Chandrasekaran et al., 2009; Ulanovsky, Las, & Nelken, 2003). Here, we document a musician advantage for neurally representing the  $F_0$ , an acoustic correlate of pitch, in a predictable context and show that the extent of this repetition-induced enhancement relates to SIN perceptual ability. These findings highlight the possibility that musicians' abilities to detect and lock on to regularities within an auditory environment contribute to their enhanced ability to perceive speech in noise.

### 3.4. Future directions

In this study, we explored the effect of stimulus regularity on the neural processing of an isolated speech sound. While we interpret our results as indicative of a greater impact of stimulus regularity on the neural encoding of speech in musicians relative to nonmusicians, future work using additional speech sounds and words will aid in the generalization of these findings. It will also be important to determine whether these effects extend to other acoustic domains, such as to music. Given that musicians demonstrate enhancements in the neural processing of pitch in both speech and music (Bidelman et al., 2011; Magne et al., 2006; Moreno et al., 2009; Musacchia et al., 2007), exploration of the impact of stimulus

regularity on the neural encoding of speech and music may provide a powerful approach for elucidating the extent to which cross-domain neural functions undergird music and language processing.

While the detection of acoustic regularities is an important aspect of auditory scene analysis, the ability to detect unexpected novel stimuli is also necessary for an individual to respond to behaviourally relevant changes in the environment. In addition to tracking sensory regularities, the brain also has the capacity to act as a novelty detector (e.g., the mismatch negativity response, or MMN (Näätänen, Gaillard, & Mäntysalo, 1978)). We propose that regularity and novelty detection within an ongoing sound stream may not be distinct processes; rather, they reflect different faces of the same coin. For example, a sound can only be deemed novel if it violates predictions generated by identified regularities. In this study, we document enhanced neural encoding of a predictably relative to a variably presented speech sound. Future work employing a more diverse range of speech sound probabilities is needed to more explicitly define the effect of statistical regularity on subcortical response properties. Further work might also vary the behavioural relevance of eliciting stimuli (e.g., behaviourally relevant regularities in language versus irrelevant regularities occurring in a steady hum of traffic, enabling it to be easily ignored). This may permit the definition of how behavioural relevance alters the neural encoding of acoustic regularities.

Our documentation of subcortical enhancement to a repeated and, hence, predictable speech sound might appear at odds with the well-documented decrease in cortical evoked responses in which stimulus repetition leads to reduced neural response magnitudes (i.e., repetition suppression). Despite its widespread acknowledgement, the neural mechanisms underlying repetition suppression remain debated. Although some have proposed decreased neuronal activity with stimulus repetition, repetition may also engender a more precise neural representation of the evoking stimulus, facilitating certain aspects of the neural response while inhibiting others (e.g., more precise inhibitory sidebands surrounding a facilitated response to the physical dimensions of the stimulus) (Grill-Spector, Henson, & Martin, 2006). Indeed, stimulus repetition does not always lead to reduced cortical evoked response magnitudes; there are certain circumstances in which repetition (Loveless, Hari, Hämäläinen, & Tiihonen, 1989; Loveless, Levänen, Jousmäki, Sams, & Hari, 1996; Wang, Mouraux, Liang, & Iannetti, 2008) and stimulus predictability (Bendixen et al., 2009) result in greater cortical magnitudes. Subcortical responses also demonstrate this duality of enhancement (Chandrasekaran et al., 2009; Dean et al., 2005; Skoe & Kraus, 2010b) or attenuation (Anderson, Christianson, & Linden, 2009; Malmierca et al., 2009; Perez-Gonzalez et al., 2005) to repeated or predictable sounds. Although our results indicate enhanced subcortical responses with stimulus repetition, especially in musicians, whether this corresponds with cortical response suppression or augmentation is undetermined. Future work employing this experimental paradigm with the concurrent recording of subcortical and cortical responses may advance our understanding of the effect of repetition and predictability on the auditory system as a whole.

#### 4. Conclusions

In order to deal with the enormous amount of sensory information received on a moment-to-moment basis, the brain must decide which elements are most relevant to the task in hand. To accomplish this goal, the nervous system continually generates predictions based on previous experiences to more efficiently extract the most relevant cues from the complex sensory environment. This adaptability is keenly displayed while attempting to hear in noise, when the auditory system's goal is to segregate the target

voice from the background noise. Pitch provides a good candidate for facilitating this process, as it can aid in distinguishing a specific voice from background noise. Here, we demonstrate that musicians are more sensitive to regularities within an ongoing speech stream, resulting in a greater neural representation of the fundamental frequency in a predictable compared to a variable context. In musicians, this increase in neural sensitivity to speech in a predictable context relates to SIN perception, thereby offering a neural basis for musicians' behavioural advantage for hearing in noise. Furthermore, the extent of musical training relates to the degree of regularity-induced neural enhancement, suggesting that this is a malleable process that can be shaped by experience.

#### Acknowledgements

This work was funded by the National Science Foundation (SGER 0842376). The authors wish to thank Jane Hornickel for her assistance with data processing.

#### References

- Ahissar, M. & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8, 457–464.
- Ahissar, M., Lubin, Y., Putter-Katz, H. & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience*, 9, 1558–1564.
- Ahissar, M., Nahum, M., Nelken, I. & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 364, 285–299.
- Aiken, S. J. & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, 245, 35–47.
- Akhoun, I., Gallégo, S., Moulin, A., Ménard, M., Veuillet, E., Berger-Vachon, C., et al. (2008). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clinical Neurophysiology*, 119, 922–933.
- Anderson, L. A., Christianson, G. B. & Linden, J. F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *The Journal of Neuroscience*, 29, 7359.
- 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.
- Assmann, P. F. & Summerfield, Q. (1987). Perceptual segregation of concurrent vowels. *Journal of the Acoustical Society of America*, 82, S120.
- 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, 253–260.
- Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, 10, 93–94.
- Baumann, O. & Belin, P. (2010). Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychological Research*, 74, 110–120.
- Bench, J., Kowal, A. & Bamford, J. (1979). The BKB (Bamford–Kowal–Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13, 108–112.
- Bendixen, A., Roeber, U. & Schröger, E. (2007). Regularity extraction and application in dynamic auditory stimulus sequences. *Journal of Cognitive Neuroscience*, 19, 1664–1677.
- Bendixen, A., Schroger, E. & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. *Journal of Neuroscience*, 29, 8447.
- Besson, M. & Ffytche, D. H. (1995). An Event-Related Potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278–1296.
- Besson, M., Ffytche, D. H. & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience Letters*, 168, 101–105.
- Bidelman, G. M., Gandour, J. T. & Krishnan, A. (2011). 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.
- Bird, J. & Darwin, C. J. (1998). Effects of a difference in fundamental frequency in separating two sentences. In A. R. Almer, A. Rees, A. Q. Summerfield, & R. Meddis (Eds.), *Psychophysical and physiological advances in hearing* (pp. 263–269). London: Whurr.
- Brattico, E., Näätänen, R. & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception*, 19, 199–222.
- Brox, J. P. & Nooteboom, S. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, 10, 23–26.
- Brown, L., Sherbenou, R. & Johnsen, S. (1982). *Test of nonverbal intelligence: TONI*. Pro-ed.
- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T. & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron*, 64, 311–319.

- Chandrasekaran, B. & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins. *Psychophysiology*, 47, 236–246.
- Clarke, F. & Becker, R. (1969). Comparison of techniques for discriminating among talkers. *Journal of Speech and Hearing Research*, 12, 747–761.
- Culling, J. F. & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping by  $F_0$ . *Journal of the Acoustical Society of America*, 93, 3454–3467.
- Darwin, C. J. (2005). Pitch and auditory grouping. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception*. Springer Handbook of Auditory Research.
- Dean, I., Harper, N. S. & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, 8, 1684–1689.
- Dean, I., Robinson, B. L., Harper, N. S. & McAlpine, D. (2008). Rapid neural adaptation to sound level statistics. *Journal of Neuroscience*, 28, 6430–6438.
- Drake, C., Jones, M. & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, 77, 251–288.
- Drullman, R. & Bronkhorst, A. (2004). Speech perception and talker segregation: Effects of level, pitch, and tactile support with multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, 116, 3090–3098.
- Engel, A., Fries, P. & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2, 704–716.
- Evans, J., Saffran, J. & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52, 321.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 360, 815–836.
- Fritz, J. B., Elhilali, M. & Shamma, S. A. (2005). Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *Journal of Neuroscience*, 25, 7623–7635.
- Fujioka, T., Trainor, L., Ross, B., Kakigi, R. & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, 16, 1010–1021.
- Fujioka, T., Trainor, L., Ross, B., Kakigi, R. & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, 17, 1578–1592.
- Galbraith, G., Arbagey, P., Branski, R., Comerchi, N. & Rector, P. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport*, 6, 2363–2367.
- Gao, E. & Suga, N. (1998). Experience-dependent corticofugal adjustment of mid-brain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 12663–12670.
- Gaser, C. & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, 23, 9240–9245.
- Gorga, M., Abbas, P. & Worthington, D. (1985). *Stimulus calibration in ABR measurements*. San Diego, CA: College-Hill Press.
- Grill-Spector, K., Henson, R. & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10, 14–23.
- Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzeliier, J. H. & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience*, 25, 10494–10501.
- Hannon, E. E., Snyder, J. S., Eerola, T. & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 956–974.
- Hutchinson, S., Lee, L. H., Gaab, N. & Schlaug, G. (2003). Cerebellar volume of musicians. *Cerebral Cortex*, 13, 943–949.
- Jentschke, S. & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, 47, 735–744.
- Koelsch, S., Jentschke, S., Sammler, D. & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, 44, 476–490.
- Kraus, N. & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11, 599–605.
- Kreiman, J., Gerratt, B., Precoda, K. & Berke, G. (1992). Individual differences in voice quality perception. *Journal of Speech and Hearing Research*, 35, 512.
- Krishnan, A. & Gandour, J. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, 110, 135–148.
- Krishnan, A., Xu, Y., Gandour, J. & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25, 161–168.
- Krumhansl, C. L. (1985). Perceiving tonal structure in music. *American Scientist*, 73, 371–378.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Kujala, T., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J. & Näätänen, R. (2000). Basic auditory dysfunction in dyslexia as demonstrated by brain activity measurements. *Psychophysiology*, 37, 262–266.
- Kujala, T. & Naatanen, R. (2001). The mismatch negativity in evaluating central auditory dysfunction in dyslexia. *Neuroscience & Biobehavioral Reviews*, 25, 535–543.
- Large, E. W. & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159.
- Lee, K. M., Skoe, E., Kraus, N. & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience*, 29, 5832–5840.
- Loveless, N., Hari, R., Hämäläinen, M. & Tiihonen, J. (1989). Evoked responses of human auditory cortex may be enhanced by preceding stimuli. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 74, 217–227.
- Loveless, N., Levänen, S., Jousmäki, V., Sams, M. & Hari, R. (1996). Temporal integration in auditory sensory memory: Neuromagnetic evidence. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 100, 220–228.
- Luo, F., Wang, Q., Kashani, A. & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience*, 28, 11615–11621.
- Magne, C., Schön, 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, 199–211.
- Malmierca, M. S., Cristaudo, S., Perez-Gonzalez, D. & Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *Journal of Neuroscience*, 29, 5483–5493.
- Margulis, E., Mlsna, L. M., Uppunda, A. K., Parrish, T. B. & Wong, P. C. M. (2009). Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Human Brain Mapping*, 30, 267–275.
- 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.
- Muller, J. R., Metha, A. B., Krauskopf, J. & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, 285, 1405–1408.
- 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.
- Näätänen, R., Gaillard, A. W. K. & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313–329.
- Näätänen, R. & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125, 826–859.
- Nikjeh, D., Lister, J. & Frisch, S. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear and Hearing*, 30, 432.
- 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.
- Oxenham, A. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, 12, 316–331.
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A. & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, 12, 169–174.
- 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, 14100.
- Parbery-Clark, A., Skoe, E., Lam, C. & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30, 653.
- 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, e18082.
- Perez-Gonzalez, D., Malmierca, M. S. & Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience*, 22, 2879–2885.
- Pressnitzer, D., Sayles, M., Micheyl, C. & Winter, I. M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology*, 18, 1124–1128.
- Russo, N., Nicol, T., Zecker, S., Hayes, E. & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, 156, 95–103.
- Schon, D., Magne, C. & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41, 341–349.
- Schulte-Körne, G., Deimel, W., Bartling, J. & Remschmidt, H. (1999). Pre-attentive processing of auditory patterns in dyslexic human subjects. *Neuroscience Letters*, 276, 41–44.
- Shinn-Cunningham, B. & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, 12, 283–299.
- Skoe, E. & Kraus, N. (2010a). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing*, 31, 302–324.
- Skoe, E. & Kraus, N. (2010b). Hearing it again and again: On-line subcortical plasticity in humans. *PLoS ONE*, 5, e13645.
- Song, J., Skoe, E., Banai, K. & Kraus, N. (2010). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*, doi:10.1162/jocn.2010.21556
- 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, 1892–1902.
- Strait, D. L. & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, 2, 1–10.
- Strait, D. L., Chan, K., Ashley, R., & Kraus, N. (in press). Specialization among the specialized: Auditory brainstem function is tuned in to timbre. *Cortex*. doi:10.1016/j.cortex.2011.03.015.
- Strait, D. L., Hornickel, J., & Kraus, N. (in press). Subcortical processing of speech regularities predicts reading and music aptitude in children. *Behavioral and Brain Functions*.
- 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.
- Suga, N. & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neuroscience*, *4*, 783–794.
- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K. & Näätänen, R. (1997). The musical brain: Brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, *226*, 1–4.
- 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*, 1636–1642.
- Tzounopoulos, T. & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*, 463–469.
- Ulanovsky, N., Las, L. & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*, 391–398.
- van Zuijlen, T., Sussman, E., Winkler, I., Näätänen, R. & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience*, *16*, 331–338.
- van Zuijlen, T., Sussman, E., Winkler, I., Näätänen, R. & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity – A mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research*, *23*, 270–276.
- Wang, A. L., Mouraux, A., Liang, M. & Iannetti, G. D. (2008). The enhancement of the N1 wave elicited by sensory stimuli presented at very short inter-stimulus intervals is a general feature across sensory systems. *PLoS ONE*, *3*, e3929.
- Webster, M. A., Kaping, D., Mizokami, Y. & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*, 557–561.
- Wen, B., Wang, G. I., Dean, I. & Delgutte, B. (2009). Dynamic range adaptation to sound level statistics in the auditory nerve. *Journal of Neuroscience*, *29*, 13797–13808.
- Winkler, I., Denham, S. L. & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*, 532–540.
- Winkler, I., Karmos, G. & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*, 239–252.
- 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.
- Yan, W. & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, *1*, 54–58.