



# Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch

Gavin M. Bidelman\*, Jackson T. Gandour, Ananthanarayan Krishnan

Department of Speech, Language, and Hearing Sciences, Purdue University, W. Lafayette, IN 47907, USA

## ARTICLE INFO

*Article history:*  
Accepted 19 July 2011

### Keywords:

Pitch discrimination  
Music perception  
Tone language  
Auditory evoked potentials  
Fundamental frequency-following response (FFR)  
Experience-dependent plasticity

## ABSTRACT

Behavioral and neurophysiological transfer effects from music experience to language processing are well-established but it is currently unclear whether or not linguistic expertise (e.g., speaking a tone language) benefits music-related processing and its perception. Here, we compare brainstem responses of English-speaking musicians/non-musicians and native speakers of Mandarin Chinese elicited by tuned and detuned musical chords, to determine if enhancements in subcortical processing translate to improvements in the perceptual discrimination of musical pitch. Relative to non-musicians, both musicians and Chinese had stronger brainstem representation of the defining pitches of musical sequences. In contrast, two behavioral pitch discrimination tasks revealed that neither Chinese nor non-musicians were able to discriminate subtle changes in musical pitch with the same accuracy as musicians. Pooled across all listeners, brainstem magnitudes predicted behavioral pitch discrimination performance but considering each group individually, only musicians showed connections between neural and behavioral measures. No brain-behavior correlations were found for tone language speakers or non-musicians. These findings point to a dissociation between subcortical neurophysiological processing and behavioral measures of pitch perception in Chinese listeners. We infer that sensory-level enhancement of musical pitch information yields cognitive-level perceptual benefits only when that information is behaviorally relevant to the listener.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

Pitch is a ubiquitous parameter of human communication which carries important information in both music and language (Plack, Oxenham, Fay, & Popper, 2005). In music, pitches are selected from fixed hierarchical scales and it is the relative relationships between such scale tones which largely contributes to the sense of a musical key and harmony in tonal music (Krumhansl, 1990). In comparison to music, tone languages provide a unique opportunity for investigating linguistic uses of pitch as these languages exploit variations in pitch at the syllable level to contrast word meaning (Yip, 2002). However, in contrast to music, pitch in language does not contain a hierarchical (i.e., scalar) framework; there is no “in-” or “out-of-tune”. Consequently, the perception of linguistic pitch patterns largely depends on cues related to the specific trajectory (i.e., contour) of pitch movement within a syllable

(Gandour, 1983, 1984) rather than distances between consecutive pitches, as it does in music (Dowling, 1978).

It is important to emphasize that linguistic pitch patterns differ substantially from those used in music; lexical tones are continuous and curvilinear (Gandour, 1994; Xu, 2006) whereas in music, pitches unfold in a discrete, stair-stepped manner (Burns, 1999, p. 217; Dowling, 1978). Indeed, specific training or long-term exposure in one domain entrains a listener to utilize pitch cues associated with that domain. Neurophysiological evidence from cortical brain potentials suggests, for instance, that musicians exploit interval based pitch cues (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Krohn, Brattico, Valimaki, & Tervaniemi, 2007) while tone language speakers exploit contour based cues (Chandrasekaran, Gandour, & Krishnan, 2007). Such “cue weighting” is consistent with each group’s unique listening experience and the relative importance of these dimensions to music (Burns & Ward, 1978) and lexical tone perception (Gandour, 1983), respectively. Given these considerable differences, it is unclear *a priori*, whether pitch experience in one domain would provide benefits in the other domain. That is, whether musical training could enhance language-related pitch processing (music-to-language transfer) or conversely, whether tone language experience could benefit music-related processing (language-to-music transfer).

*Abbreviations:* ANOVA, analysis of variance; C, Chinese; DL, difference limen; EEG, electroencephalogram; F0, fundamental frequency; FFR, frequency-following response; FFT, Fast Fourier Transform; M, musicians; NM, non-musicians.

\* Corresponding author. Present address: Rotman Research Institute, Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, ON, Canada M6A 2E1.

E-mail address: [gbidelman@rotman-baycrest.on.ca](mailto:gbidelman@rotman-baycrest.on.ca) (G.M. Bidelman).

Yet, a rapidly growing body of evidence suggests that brain mechanisms governing music and language processing interact (Bidelman, Gandour, & Krishnan, 2011; Koelsch et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Patel, 2008; Slevc, Rosenberg, & Patel, 2009). Such *cross-domain transfer effects* have now been extensively reported in the direction from *music to language*; musicians demonstrate perceptual enhancements in a myriad of language specific abilities including phonological processing (Anvari, Trainor, Woodside, & Levy, 2002), verbal memory (Chan, Ho, & Cheung, 1998; Franklin et al., 2008), formant and voice pitch discrimination (Bidelman & Krishnan, 2010), sensitivity to prosodic cues (Thompson, Schellenberg, & Husain, 2004), degraded speech perception (Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, Lam, & Kraus, 2009), second language proficiency (Slevc & Miyake, 2006), and lexical tone identification (Delogu, Lampis, & Olivetti Belardinelli, 2006; Delogu, Lampis, & Olivetti Belardinelli, 2010; Lee & Hung, 2008). These perceptual advantages are corroborated by electrophysiological evidence demonstrating that both cortical (Chandrasekaran, Krishnan, & Gandour, 2009; Marie, Magne, & Besson, 2010; Moreno & Besson, 2005; Pantev, Roberts, Schulz, Engeliem, & Ross, 2001; Schon, Magne, & Besson, 2004) and even subcortical (Bidelman, Gandour, et al., 2011; Musacchia, Strait, & Kraus, 2008; Wong, Skoe, Russo, Dees, & Kraus, 2007) brain circuitry tuned by long-term music experience facilitates the encoding of speech related signals. However, whether or not the reverse effect exists, that is, the ability of language experience to positively influence music processing/perception, remains an unresolved question (e.g., Schellenberg & Peretz, 2008; Schellenberg & Trehub, 2008).

Studies investigating transfer effects in the direction of *language-to-music* are scarce. Almost all have focused on putative connections between tone language experience and absolute pitch (AP) (e.g., Deutsch, Henthorn, Marvin, & Lee & Lee, 2010; Xu, 2006), i.e., the rare ability to label musical notes without any external reference. It has been noted, however, that AP is largely irrelevant to most music listening (Levitin & Rogers, 2005, p. 26). Musical tasks predominantly involve monitoring the *relative* relationships between pitches (i.e., intervallic distances) rather than labeling absolute pitch values in isolation. Of language-to-music studies that have focused on intervallic aspects of music, a behavioral study by Pfordresher and Brown (2009) indicated that tone language speakers were better able to discriminate the size (cf. height) of two-tone musical intervals relative to English-speaking non-musicians. In an auditory electrophysiological study, Bidelman et al. (2011) found that brainstem responses evoked by musical intervals were both more accurate and more robust in native speakers of Mandarin Chinese compared to English-speaking non-musicians, suggesting that long-term experience with linguistic pitch may transfer to subcortical encoding of musical pitch. Together, these studies demonstrate a potential for linguistic pitch experience to carry over into nonlinguistic (i.e., musical) domains (Bidelman, Gandour, et al., 2011, p. 430; Pfordresher & Brown, 2009, p. 1385). However, to date, no study has investigated *concurrently* the nature of the relationship between tone language listeners' neural processing and their behavioral perception of more complex musical stimuli (e.g., chords or arpeggios).

The aim of the present work is to determine the effect of domain-specific pitch experience on the neural processing and perception of musical stimuli. Specifically, we attempt to determine whether the previously observed superiority in Chinese listeners' subcortical representation of musical pitch (e.g., Bidelman, Gandour, et al., 2011) provides any benefit in music perception (e.g., Pfordresher & Brown, 2009). We bridge the gap between neurophysiological and behavioral studies of language-to-music transfer effects by examining brainstem responses in conjunction with corresponding perceptual discrimination of complex musical stimuli.

We have previously shown that musicians have more robust brainstem representation for defining features of tuned (i.e., major and minor) and detuned chordal arpeggios, and moreover, a superior ability to detect whether they are in or out of tune behaviorally (Bidelman, Krishnan, & Gandour, 2011). By comparing brainstem responses and perceptual discrimination of such chords in Chinese non-musicians vs. English-speaking musicians and non-musicians, we are able to assess not only the extent to which *linguistic pitch experience* enhances pre-attentive, subcortical encoding of *musical chords*, but also whether or not such language-dependent neural enhancements confer any behavioral advantages in music perception.

## 2. Methods

### 2.1. Participants

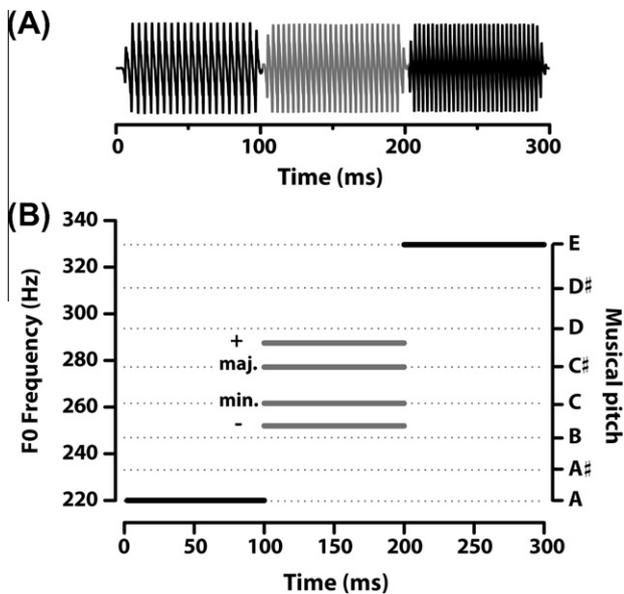
Eleven English-speaking musicians (M: 7 male, 4 female), 11 English-speaking non-musicians (NM: 6 male, 5 female), and 11 native speakers of Mandarin Chinese (C: 5 male, 6 female) were recruited from the Purdue University student body to participate in the experiment. All participants exhibited normal hearing sensitivity at audiometric frequencies between 500 and 4000 Hz and reported no previous history of neurological or psychiatric illnesses. They were closely matched in age (M:  $22.6 \pm 2.2$  yrs, NM:  $22.8 \pm 3.4$  yrs, C:  $23.7 \pm 4.1$  yrs), years of formal education (M:  $17.1 \pm 1.8$  yrs, NM:  $16.6 \pm 2.6$  yrs, C:  $17.2 \pm 2.8$  yrs), and were strongly right handed (laterality index  $>73\%$ ) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). Musicians were amateur instrumentalists having  $\geq 10$  yrs of continuous instruction on their principal instrument ( $\mu \pm \sigma$ ;  $12.4 \pm 1.8$  yrs), beginning at or before the age of 11 ( $8.7 \pm 1.4$  yrs) (Table 1). All currently played his/her instrument(s). Non-musicians had  $\leq 1$  yr of formal music training ( $0.5 \pm 0.5$  yrs) on any combination of instruments. No English-speaking participant had any prior experience with a tone language. Chinese participants were born and raised in mainland China. All Chinese subjects were classified as late onset Mandarin/English bilinguals who exhibited moderate proficiency in English, as determined by a language history questionnaire (Ping, Sepanski, & Zhao, 2006). They used their native language in the majority ( $M = 67\%$ ) of their combined daily activities. None had received formal instruction in English before the age of 9 ( $12.5 \pm 2.7$  yrs) nor had more than 3 yrs of musical training ( $0.8 \pm 1.1$  yrs). Each participant was paid for his/her time and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

### 2.2. Stimuli

Four triad arpeggios (i.e., three-note chords played sequentially) were constructed which differed only in their chordal third (Fig. 1;

**Table 1**  
Musical background of musician participants.

Participant	Instrument(s)	Years of training	Age of onset
M1	Trumpet/piano	14	10
M2	Saxophone/piano	13	8
M3	Piano/guitar	10	9
M4	Saxophone/clarinet	13	11
M5	Piano/saxophone	11	8
M6	Violin/piano	11	8
M7	Trumpet	11	9
M8	String bass	12	8
M9	Trombone/tuba	11	7
M10	Bassoon/piano	16	7
M11	Saxophone/piano	14	11
Mean (SD)	12.4 (1.8)	8.7 (1.4)	



**Fig. 1.** Triad arpeggios used to evoke brainstem responses. (A) Four sequences were created by concatenating three 100 ms pitches together (B) whose F0s corresponded to either prototypical (major, minor) or mistuned (detuned up, +; detuned down, -) versions of musical chords. Individual notes were synthesized using a tone-complex consisting of six harmonics (amplitudes =  $1/N$ , where  $N$  is the harmonic number) added in sine phase. Only the pitch of the chordal third differed between arpeggios as represented by the grayed portion of the time-waveforms (A) and F0 tracks (B). The F0 of the chordal third varied according to the stimulus: major = 277 Hz, minor = 262 Hz, detuned up = 287 Hz, detuned down = 252 Hz. Detuned thirds represent a 4% difference in F0 from the actual major or minor third, respectively. F0, fundamental frequency.

Supplemental material). Two sequences were exemplary arpeggios of Western music practice (major and minor); the other two represented detuned versions of these chords (detuned up, +; detuned down, -) whose third was slightly sharp or flat of the actual major or minor third, respectively. Individual notes were synthesized using a tone-complex consisting of six harmonics (sine phase) with 100 ms duration (including a 5 ms rise-fall time). For each sequence, the three notes were concatenated to create a contiguous chordal arpeggio 300 ms in duration (Fig. 1A). The fundamental frequency (F0) of each of the three notes (i.e., chordal root, third, fifth) per triad were as follows (Fig. 1B): major = 220, 277, 330 Hz; minor = 220, 262, 330 Hz; detuned up = 220, 287, 330 Hz; detuned down = 220, 252, 330 Hz (Fig. 1B). In the detuned arpeggios, mistuning in the chord's third represent a +4% or -4% difference in F0 from the actual major or minor third, respectively (cf. a full musical semitone which represents a 6% difference in frequency). Thus, because F0s of the first (root) and third (fifth) notes were identical across stimuli (220 and 330 Hz, respectively) the chords differed only in the F0 of their 2nd note (third).

### 2.3. Neurophysiological brainstem responses

#### 2.3.1. FFR data acquisition

As a window into the early stages of subcortical pitch processing we utilize the scalp recorded frequency-following response (FFR). The FFR reflects sustained phase-locked activity from a population of neural elements within the rostral midbrain (for review, see Krishnan, 2007) which provides a robust index of the brainstem's transcription of speech (for reviews, see Johnson, Nicol, & Kraus, 2005; Krishnan & Gandour, 2009; Skoe & Kraus, 2010) and musically relevant features (Bidelman, Gandour, et al., 2011; Bidelman & Krishnan, 2009; Bidelman & Krishnan, 2011; Bidelman, Krishnan, et al., 2011) of the acoustic signal. The FFR recording protocol was similar to that used in previous reports from our laboratory (e.g.,

Bidelman & Krishnan, 2009; Bidelman & Krishnan, 2010). Participants reclined comfortably in an acoustically and electrically shielded booth to facilitate recording of brainstem responses. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts), ignore the sounds they hear, and were allowed to sleep throughout the duration of FFR recording. FFRs were elicited from each participant by monaural stimulation of the right ear at a level of 80 dB SPL through a magnetically shielded insert earphone (ER-3A; Etymotic Research, Elk Grove Village, IL, USA). Each stimulus was presented using rarefaction polarity at a repetition rate of 2.44/s. Presentation order was randomized both within and across participants. Control of the experimental protocol was accomplished by a signal generation and data acquisition system (Intelligent Hearing Systems; Miami, FL, USA).

FFRs were obtained using a vertical electrode montage which provides the optimal configuration for recording brainstem activity (Galbraith et al., 2000). Ag-AgCl scalp electrodes placed on the midline of the forehead at the hairline ( $\sim$ Fz; non-inverting, active) and right mastoid (A2; inverting, reference) served as the inputs to a differential amplifier. Another electrode placed on the mid-forehead (Fpz) served as the common ground. The raw EEG was amplified by 200,000 and filtered online between 30 and 5000 Hz. All inter-electrode impedances were maintained  $\leq 1$  k $\Omega$ . Individual sweeps were recorded using an analysis window of 320 ms at a sampling rate of 10 kHz. Sweeps containing activity exceeding  $\pm 35$   $\mu$ V were rejected as artifacts and excluded from the final average. FFR response waveforms were further band-pass filtered offline from 100 to 2500 Hz (-6 dB/octave roll-off) to minimize low-frequency physiologic noise and limit the inclusion of cortical activity (Musacchia et al., 2008). In total, each FFR response waveform represents the average of 3000 artifact-free stimulus presentations.

#### 2.3.2. FFR data analysis

FFR encoding of information relevant to the pitch of the chord defining note (i.e., the third) was quantified by measuring the magnitude of the F0 component from the corresponding portion of the response waveform per melodic triad. FFRs were segmented into three 100 ms sections (15–115 ms; 115–215 ms; 215–315 ms) corresponding to the sustained portions of the response to each musical note. The spectrum of the second segment, corresponding to the response to the chord's third, was computed by taking the Fast Fourier Transform (FFT) of a time-windowed version of its temporal waveform (Gaussian window). For each subject, the magnitude of F0 was measured as the peak in the FFT, relative to the noise floor, which fell in the same frequency range expected by the input stimulus (note 2: 245–300 Hz; see stimulus F0 tracks, Fig. 1B). All FFR data analyses were performed using custom routines coded in MATLAB<sup>®</sup> 7.10 (The MathWorks, Inc., Natick, MA, USA).

### 2.4. Behavioral measures of musical pitch discrimination

#### 2.4.1. F0 discrimination

Behavioral fundamental frequency difference limens (F0 DLs) were measured for each participant in a two-alternative forced choice (2AFC) discrimination task (e.g., Bidelman & Krishnan, 2010; Krishnan, Bidelman, & Gandour, 2010). For a given trial, participants heard two sequential intervals, one containing a reference pitch (i.e., tone complex) and one containing a comparison, assigned randomly. The reference pitch had a fixed F0 frequency of 270 Hz (the average F0 of all chordal stimuli, see Fig. 1); the pitch of the comparison was always greater (i.e., higher F0). The participant's task was to identify the interval which contained a higher sounding pitch. Following a brief training run, discrimination thresholds were measured using a two-down, one-up adaptive paradigm which tracks the 71% correct point on the psychometric function (Levitt, 1971). Following two consecutive correct

responses, the frequency difference of the comparison was decreased for the subsequent trial, and increased following a single incorrect response. Frequency difference between reference and comparison intervals was varied using a geometric step size of  $\sqrt{2}$  between response reversals. Sixteen reversals were measured and the geometric mean of the last 12 taken as the individual's F0 DL, that is, the minimum frequency difference needed to detect a change in pitch.

#### 2.4.2. Chordal detuning discrimination

Five participants from each group took part in a second pitch discrimination task which was conducted to more closely mimic the FFR protocol and determine whether groups differed in their ability to detect chordal detuning at a perceptual level. Discrimination sensitivity was measured separately for the three most meaningful stimulus pairings (major/minor, major/detuned up, minor/detuned down) using a same-different task (Bidelman, Krishnan, et al., 2011). For each of these three conditions, participants heard 100 pairs of the chordal arpeggios presented with an interstimulus interval of 500 ms. Half of these trials contained chords with different thirds (e.g., major-detuned up) and half were catch trials containing the same chord (e.g., major-major), assigned randomly. After hearing each pair, participants were instructed to judge whether the two chord sequences were the “same” or “different” via a button press on the computer. The number of hits and false alarms were recorded per condition. Hits were defined as “different” responses to a pair of physically different stimuli and false alarms as “different” responses to a pair in which the items were actually identical. All stimuli were presented at  $\sim 75$  dB SPL through circumaural headphones (Sennheiser HD 580; Sennheiser Electronic Corp., Old Lyme, CT, USA). Stimulus presentation and response collection for both behavioral tasks were implemented in custom interfaces coded in MATLAB.

#### 2.5. Statistical analysis

A two-way, mixed-model ANOVA (SAS<sup>®</sup>; SAS Institute, Inc., Cary, NC, USA) was conducted on F0 magnitudes derived from FFRs in order to evaluate the effects of pitch experience (i.e., musical, linguistic, none) and stimulus context (i.e., prototypical vs. non-prototypical sequence) on brainstem encoding of musical pitch. Group (three levels; musicians, Chinese, non-musicians) functioned as the between-subjects factor and stimulus (4 levels; major, minor, detuned up, detuned down) as the within-subjects factor. An *a priori* level of significance was set at  $\alpha = 0.05$ . All multiple pairwise comparisons were adjusted with Bonferroni corrections. Where appropriate, partial eta-squared ( $\eta_p^2$ ) values are reported to indicate effect sizes.

Behavioral discrimination sensitivity scores ( $d'$ ) were computed using hit ( $H$ ) and false alarm ( $FA$ ) rates (i.e.,  $d' = z(H) - z(FA)$ , where  $z(\cdot)$  represents the z-score operator). Based on initial diagnostics and the Box-Cox procedure (Box & Cox, 1964), both  $d'$  and F0 DL scores were log-transformed to improve normality and homogeneity of variance assumptions necessary for a parametric ANOVA. Log-transformed  $d'$  scores were then submitted to a two-way mixed model with group as the between-subjects factor and stimulus pair (three levels; major/detuned up, minor/detuned down, major/minor) as the within-subjects factor. Transformed F0 DLs were analyzed using a similar one-way model with group as the between-subjects factor of interest.

### 3. Results

FFR time-waveforms in response to the major chord are shown for the three groups in Fig. 2. Similar time-waveforms were

observed in response to the other three stimuli. For all groups, clear onset components (i.e., large negative deflections) are seen at the three time marks corresponding to the individual onset of each note (note 1:  $\sim 17$  ms, note 2:  $\sim 117$  ms, note 3:  $\sim 217$  ms). Relative to non-musicians, musician and Chinese FFRs contain larger amplitudes during the chordal third (2nd note,  $\sim 110$  to 210 ms), the defining pitch of the sequence. Within this same time window, non-musician responses show reduced amplitude indicating poorer representation of this chord-defining pitch (see also Fig. 3A).

#### 3.1. Neural representation of chordal thirds

FFR encoding of F0 for the thirds of chordal standard and detuned arpeggios are shown in Fig. 3. Individual panels show the meaningful comparisons contrasting a prototypical musical chord with its detuned counterpart: panel A, major vs. detuned up; panel B, minor vs. detuned down. An omnibus ANOVA on F0 encoding revealed significant main effects of group [ $F_{2,30} = 14.01$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.48$ ] and stimulus [ $F_{3,90} = 8.30$ ,  $p = 0.0001$ ,  $\eta_p^2 = 0.22$ ] on F0 encoding, as well as a group  $\times$  stimulus interaction [ $F_{6,90} = 2.89$ ,  $p = 0.0126$ ,  $\eta_p^2 = 0.16$ ].

*A priori* contrasts revealed that regardless of the eliciting arpeggio, musicians' brainstem responses contained a larger F0 magnitude than non-musicians. Chinese, on the other hand, had larger F0 magnitudes than non-musicians only for the detuned stimuli. No differences were found between musicians and Chinese for any of the four triads. These group comparisons ( $[M = C] > NM$ ) indicate that either music or linguistic pitch experience is mutually beneficial to brainstem mechanisms implicated in music processing.

By group, F0 magnitude did not differ across triads for either the musicians or Chinese, indicating superior encoding in both cohorts

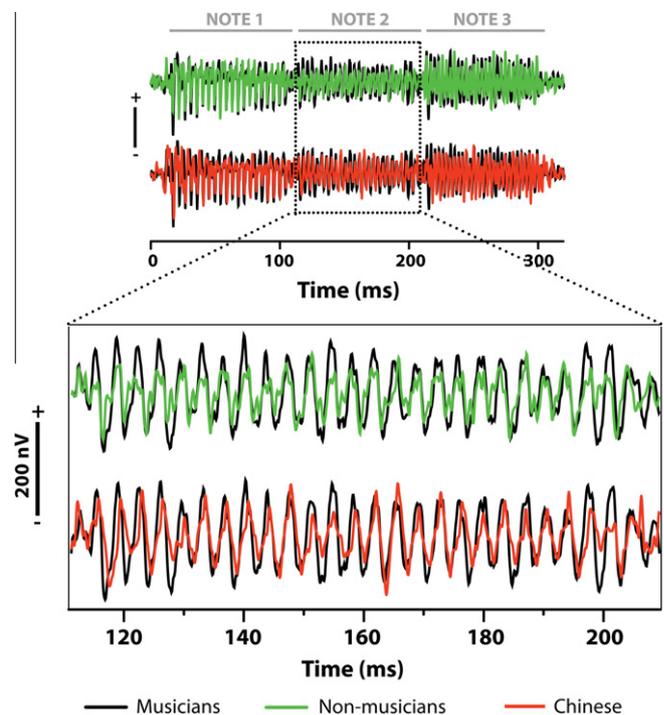
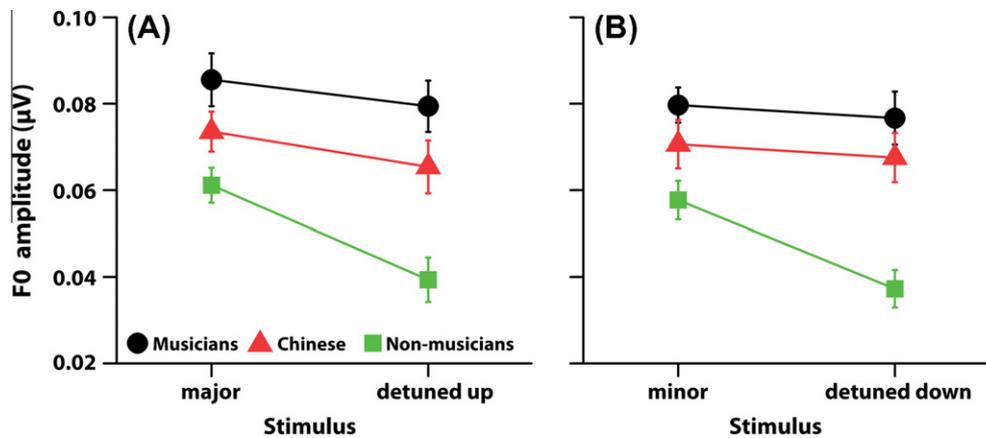


Fig. 2. Representative FFR time-waveforms evoked by the major triad. To ease group comparisons, Chinese and non-musician responses are overlaid on those of the musician group. As shown by the expanded inset, musicians show larger amplitudes than non-musician listeners during the time window of the chordal third (i.e., 2nd note), the defining pitch of the sequence. Musicians and Chinese show little difference in the same time window. Scale bars = 200 nV.



**Fig. 3.** Group comparisons of brainstem pitch encoding for the defining note of musical triads (i.e., the third). Relative to non-musicians, musicians and Chinese show enhanced F0 magnitudes in response to both prototypical arpeggios and those which are slightly sharp (A) or flat (B) of standard musical chords. Yet, no differences in FFR are found between musicians and Chinese listeners on any of the stimuli. Thus, musical and linguistic pitch experience provides mutual enhancements to brainstem representations of in- and out-of tune musical chords. When the third of the chord is slightly sharp (+4%) or flat (−4%) relative to the major and minor third, respectively, both musicians and Chinese encode the pitch of detuned notes equally as well as tempered notes (e.g., panel A, compare F0 magnitudes between major and detuned up). Non-musicians, on the other hand, show a marked decrease in F0 magnitude when the chord is detuned from the standard major or minor prototype. Musician and non-musician data are re-plotted from Bidelman et al. (2011). F0, fundamental frequency.

regardless of whether the chordal third was major or minor, in or out of tune, and irrespective of the listener's domain of pitch expertise. For non-musicians, F0 encoding was identical between the major and minor chords, two of the most regularly occurring sequences in music (Budge, 1943). However, it was significantly reduced for the detuned up and down sequences in comparison to the major and minor chords, respectively. Together, these results indicate brainstem encoding of musical pitch is impervious to changes in chordal temperament for musicians and Chinese but is diminished with chordal detuning in non-musicians.

## 3.2. Behavioral measures of musical pitch discrimination

### 3.2.1. F0 discrimination performance

Mean F0 DLs are shown per group in Fig. 4A. An ANOVA revealed a significant main effect of group on F0 DLs [ $F_{2,30} = 19.90$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.57$ ]. Post hoc multiple comparisons revealed that musicians obtained significantly better (i.e., smaller) DLs than musically naïve subjects ( $M < [C = NM]$ ), meaning that musicians were better able to detect minute changes in absolute pitch. On average, musicians' DLs were approximately 3–3.5 times smaller than those of the musically untrained subjects (C and NM).

### 3.2.2. Chordal discrimination performance

Group behavioral chordal discrimination sensitivity, as measured by  $d'$ , are shown in Fig. 4B. Values represent one's ability to discriminate melodic triads where only the third of the chord differed between stimulus pairs. By convention,  $d' = 1$  (dashed line) represents performance threshold and  $d' = 0$ , chance performance. An ANOVA on  $d'$  scores revealed significant main effects of group [ $F_{2,12} = 13.90$ ,  $p = 0.0008$ ,  $\eta_p^2 = 0.70$ ] and stimulus pair [ $F_{2,24} = 16.21$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.57$ ], as well as a group  $\times$  stimulus pair interaction [ $F_{4,24} = 3.92$ ,  $p = 0.0137$ ,  $\eta_p^2 = 0.40$ ]. By group, *a priori* contrasts revealed that musicians' ability to discriminate melodic triads was well above threshold regardless of stimulus pair. No significant differences in their discrimination ability were observed between standard (major/minor) and detuned (major/up, minor/down) stimulus pairs. Chinese and non-musicians, on the other hand, were able to achieve above threshold performance only when discriminating the major/minor pair. They could not accurately distinguish the detuned stimulus pairs (major/up, minor/down). These results indicate that only musicians perceive minute

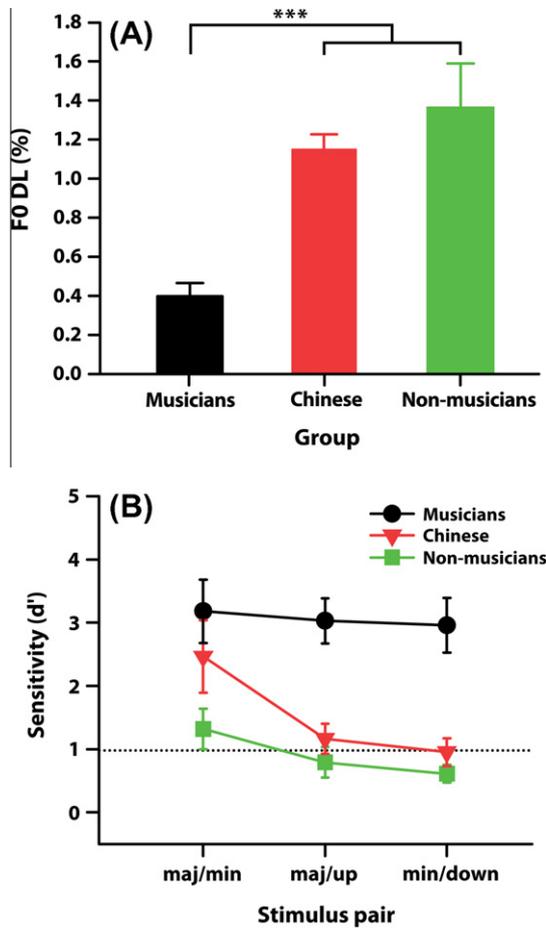
changes in musical pitch, which are otherwise undetectable by musically untrained listeners (C, NM).

By condition, musicians' ability to discriminate melodic triads was superior to non-musicians across all stimulus pairs. Importantly, musicians were more accurate in discriminating detuned stimulus pairs (major/up, minor/down) than Chinese listeners but did not differ in their major/minor discrimination. No significant differences in chordal discrimination were observed between Chinese and non-musicians.

## 3.3. Connections between neurophysiological responses to musical pitch

To investigate brain-behavior connections (or lack thereof), we regressed neural (FFR) and behavioral (F0 DL/chordal discrimination) measures against one another to determine the extent to which subcortical responses to detuned musical chords could predict perceptual performance in both behavioral tasks. Note that only  $n = 5$  subjects per group participated in the chordal discrimination task. Pearson's correlations ( $r$ ) were computed between brainstem F0 magnitudes elicited in each of the four stimulus conditions and behavioral F0 DLs. Pooled across all listeners, only brainstem responses in the detuned up condition showed close correspondence with F0 DL performance ( $r_{all} = 0.57$ ,  $p < 0.001$ ). By group, only musicians revealed a connection between neural and perceptual measures ( $r_M = 0.70$ ,  $p = 0.01$ ); neither Chinese ( $r_C = 0.19$ ,  $p = 0.57$ ) nor non-musicians ( $r_{NM} = 0.34$ ,  $p = 0.30$ ) exhibited a brain-behavior correlation for this same condition (Fig. 5A). A significant association between these measures in musicians suggests that better pitch discrimination performance (i.e., smaller F0 DLs) is, at least in part, predicted by more robust FFR encoding. The fact that a FFR/F0 DL correlation is found only in musicians indicates that such a brain-behavior connection is limited to musical training and not the result of pitch experience per se (e.g., no correlation is found in the Chinese group).

Across groups, brainstem F0 encoding in the detuned up condition also predicted listeners' performance in discriminating the major chord from its detuned version, as measured by  $d'$  ( $r = 0.76$ ,  $p = 0.001$ ; Fig. 5B). Note that a similar correspondence ( $r = 0.60$ ,  $p = 0.018$ ) was observed between F0 encoding in the detuned down condition and the corresponding condition in the behavioral task (min/down; data not shown). In this neural-perceptual space,

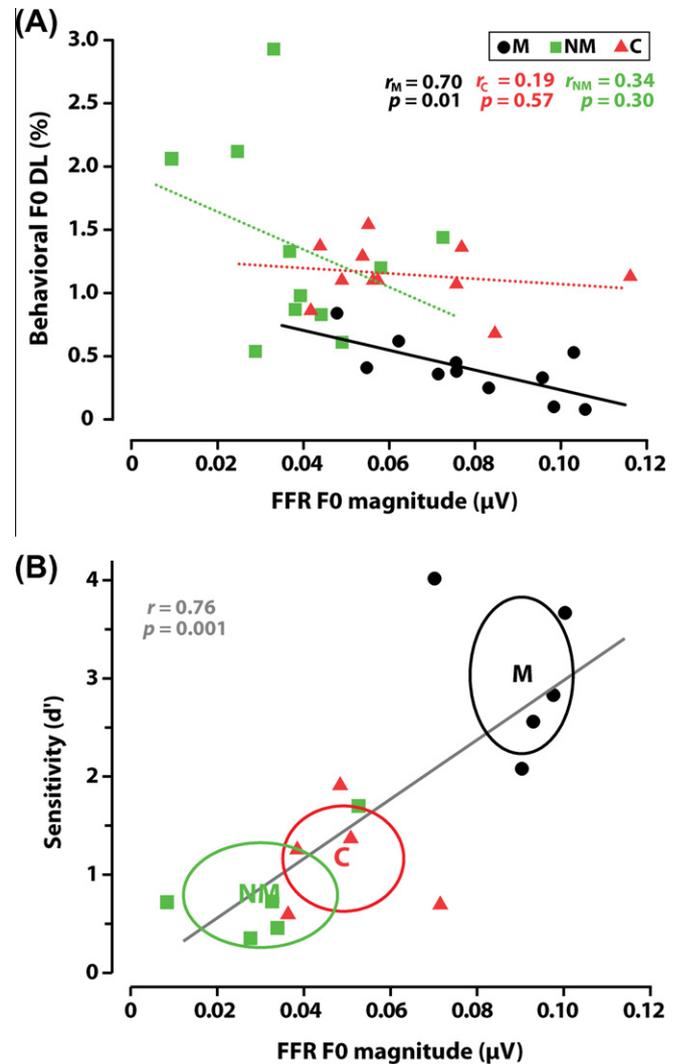


**Fig. 4.** Perceptual benefits for musical pitch discrimination are limited to musicians. (A) Group comparisons for behavioral frequency difference limens (F0 DLs). Musicians' pitch discrimination thresholds are  $\sim 3$  times smaller (i.e., better) than either the English-speaking non-musician or Chinese group, whose performance did not differ from one another. The standard F0 was 270 Hz. (B) Group  $d'$  scores for discriminating chord arpeggios ( $n = 5$  per group). By convention, discrimination threshold is represented by a  $d' = 1$  (dashed line). Musicians discriminate all chord pairings well above threshold, including standard chords (major/minor) as well as sequences in which the third is out of tune (major/up, minor/down). In contrast, non-musicians and Chinese discriminate only the major/minor pair above threshold. They are unable to accurately distinguish standard from detuned sequences (major/up, minor/down). Musician and non-musician data are re-plotted from Bidelman et al. (2011). Error bars =  $\pm 1$  SE; \*\*\* $p < 0.001$ .

musicians appear maximally separated from musically untrained listeners (Chinese and non-musicians); they have more robust neurophysiological encoding for musical pitch and maintain a higher fidelity (i.e., larger  $d'$ ) in perceiving minute deviations in chordal arpeggios. The correspondence between FFR encoding for out of tune chords and behavioral detection of such detuning across listeners suggests that an individual's ability to distinguish pitch deviations in musical sequences is, in part, predicted by how well such auditory features are encoded at a subcortical level.

#### 4. Discussion

By measuring brainstem responses to prototypical and detuned musical chord sequences we found that listeners with extensive pitch experience (musicians and tone-language speakers) have enhanced subcortical representation for musically relevant pitch when compared to inexperienced listeners (English speaking non-musicians). Yet, despite the relatively strong brainstem encoding in the Chinese (neurophysiological language-to-music



**Fig. 5.** (Dis)associations between brainstem encoding and perceptual discrimination of musical pitch. (A) F0 magnitudes computed from brainstem responses to chordal detuning (up condition) predict behavioral discrimination for musicians ( $r_M = 0.70$ ,  $p = 0.01$ ). That is, better pitch discrimination (i.e., smaller F0 DLs) is predicted by more robust FFR encoding. In contrast to musicians, no correspondence exists between brainstem measures and F0 DLs for Chinese or non-musician listeners (dotted lines,  $p > 0.05$ ). (B) Across groups, brainstem F0 encoding predicts a listener's ability to distinguish in and out of tune chords (+4% change in F0), i.e., more robust FFR magnitudes correspond to higher perceptual sensitivity ( $d'$ ). Each point represents an individual listener's FFR measure for the detuned up condition plotted against his/her performance in the maj/up discrimination task ( $n = 5$  per group). The centroid of each ellipse gives the grand average for each group while its radius denotes  $\pm 1$  SD in the neural and behavioral dimension, respectively. Note the systematic clustering of groups and musicians' maximal separation from musically untrained listeners (C and NM) in the neural-perceptual space.

transfer effect), behavioral measures of pitch discrimination and chordal detuning sensitivity reveal that this neural improvement does not necessarily translate to perceptual benefits as it does in the musician group. Indeed, tone language speakers performed no better than (i.e., as poorly as) English non-musicians in discriminating musical pitch stimuli.

##### 4.1. Cross-domain effects in brainstem encoding of musical chords

Our findings provide further evidence for experience-dependent plasticity induced by long-term experience with pitch. Both musicians and Chinese exhibit stronger pitch encoding of three-note chords as compared to individuals who are untrained musically

and lack any exposure to tonal languages (non-musicians) (Figs. 2 and 3). Their FFRs show no appreciable reduction in neural representation of pitch with parametric manipulation of the chordal third (i.e., major or minor, in or out of tune). These findings demonstrate that sustained brainstem activity is enhanced after long-term experience with pitch regardless of a listener's specific domain of expertise (i.e., speaking Mandarin vs. playing an instrument). They also converge with previous studies which demonstrate that subcortical pitch processing is not hardwired, but rather, is malleable and depends on an individual's auditory history and/or training (Bidelman, Gandour, et al., 2011; Bidelman & Krishnan, 2010; Carcagno & Plack, 2011; Krishnan, Gandour, & Bidelman, 2010; Krishnan, Gandour, Bidelman, & Swaminathan, 2009; Musacchia et al., 2008; Parbery-Clark, Skoe, & Kraus, 2009).

It is possible that the brainstem response enhancements we observe in our "pitch experts" originate as the result of top-down influence. Extensive descending projections from cortex to the inferior colliculus (IC) which comprise the corticofugal pathway are known to modulate/tune response properties in the IC (Suga, 2008), the presumed neural generator of the FFR in humans (Sohmer, Pratt, & Kinarti, 1977). Indeed, animal data indicates that corticofugal feedback strengthens IC response properties as a stimulus becomes behaviorally relevant to the animal through associative learning (Gao & Suga, 1998). Thus, it is plausible that during tone language or music acquisition, corticofugal influence may act to strengthen and hence provide a response "gain" for behaviorally relevant sound (e.g., linguistic tones, musical melodies). Interestingly, this experience-dependent mechanism seems somewhat general in that both tone language and musical experience lead to enhancements in brainstem encoding of pitch, regardless of whether the signal is specific to music or language per se (present study; Bidelman, Gandour, et al., 2011).

In contrast to musicians and Chinese listeners, brainstem responses of non-musicians are differentially affected by the musical context of arpeggios. This context effect results in diminished magnitudes for detuned chords relative to their major/minor counterparts (Fig. 3). The more favorable encoding of prototypical musical sequences may be likened to the fact that even non-musicians are experienced listeners with the major and minor triads (e.g., Bowling, Gill, Choi, Prinz, & Purves, 2010) which are among the most commonly occurring chords in tonal music (Budge, 1943; Eberlein, 1994). Indeed, based on neural data alone, it would appear *prima facie* that non-musicians encode only veridical musical pitch relationships and "filter out" detunings. The fact that their neural responses distinguish prototypical from non-prototypical musical arpeggios notwithstanding (Fig. 3), such differentiation appears to play no cognitive role. Non-musicians show the poorest behavioral performance in pitch and chordal discrimination, especially when judging true musical chords from their detuned counterparts (Fig. 4B).

Despite their lack of experience with musical pitch patterns, Chinese non-musicians show superior encoding for chord-defining pitches relative to English-speaking non-musicians. Thus, the neurophysiological benefits afforded by tone language experience extend beyond the bounds of language-specific stimuli. Music-to-language transfer effects are well documented as evidenced by several studies demonstrating enhanced cortical and subcortical linguistic pitch processing in musically trained listeners (e.g., Besson, Schon, Moreno, Santos, & Magne, 2007; Bidelman, Gandour, et al., 2011; Bidelman & Krishnan, 2010; Marques, Moreno, Castro, & Besson, 2007; Moreno & Besson, 2005; Moreno et al., 2009; Parbery-Clark, Skoe, & Kraus, 2009; Schon et al., 2004; Wong et al., 2007). Our findings provide evidence for the reverse language-to-music transfer effect, i.e., neurophysiological enhancement of musical pitch processing in native speakers of a tone language (cf. Bidelman, Gandour, et al., 2011; Marie, Kujala, &

Besson, 2010). Thus, although the origin of one's pitch expertise may be exclusively musical or linguistic in nature, years of active listening to complex pitch patterns in and of themselves, regardless of domain, mutually sharpens neural mechanisms recruited for both musical and linguistic pitch processing.

#### 4.2. Behavioral benefits for musical pitch perception are limited to musicians

In contrast to the neurophysiological findings, behaviorally, only musicians achieve superior performance in detecting deviations in pitch and chordal temperament. On average, musicians' F0 DLs were approximately 3–3.5 times smaller (i.e., better) than those of the musically untrained subjects (i.e., C and NM; Fig. 4A). These data are consistent with previous reports demonstrating superior pitch discrimination thresholds for musicians (Bidelman & Krishnan, 2010; Bidelman, Krishnan, et al., 2011; Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Strait, Kraus, Parbery-Clark, & Ashley, 2010).

Though one may suppose that linguistic pitch experience would enhance a listener's ability to detect deviations in absolute pitch (e.g., Deutsch, Henthorn, Marvin, & Xu, 2006), we observe no benefit of tone language experience on F0 DLs (Fig. 4A, C = NM). Our data converge with previous studies which have failed to demonstrate any advantage in absolute pitch sensitivity (i.e., improved JND) for tone language speakers (Bent, Bradlow, & Wright, 2006; Burns & Sampat, 1980; Pfordresher & Brown, 2009; Stargay & Downs, 1993). Prosodic information carried in language depends on pitch contour, direction, and height cues rather than absolute variations in pitch per se (e.g., Gandour, 1983; for review, see Patel, 2008, pp. 46–47). As such, the inability of our Chinese listeners to exploit absolute pitch cues as well as musicians may reflect either the irrelevance of such cues as language or alternatively, the possibility that musical training has a larger impact than linguistic experience on an individual's perceptual ability to discriminate pitch.

Similarly, in the chordal discrimination task, only musicians are able to discriminate standard and detuned arpeggios well above threshold (Fig. 4B). This finding conflicts with a previous study demonstrating that tone language speakers better discriminate the size (i.e., height) of two-note musical intervals than English-speaking non-musicians (Pfordresher & Brown, 2009). Using three-note musical chords in the present study, non-musicians and Chinese reliably discriminate major from minor chords, in which case the difference between stimuli is a full semitone. However, their discrimination fails to rise above threshold when presented with chords differing by less than a semitone (cf. C and NM on major/minor vs. major/up and minor/down, Fig. 4B). The apparent discrepancy with Pfordresher and Brown (2009) is likely attributable to differences in stimulus complexity and task demands. Whereas their two-note interval stimuli require a simple comparison of pitch height between sets of simultaneously sounding notes (harmonic discrimination), our three-note chordal arpeggios require a listener to monitor the entire triad over time, in addition to comparing the size of its constituent intervals (melodic discrimination). Thus, while tone language speakers may enjoy a perceptual advantage in simple pitch height (cf. intervallic distance) discrimination (Pfordresher & Brown, 2009), this benefit largely disappears when the listener is required to detect subtle changes (i.e., <1 semitone) *within a melodic sequence of pitches*. This ability, on the other hand, is a defining characteristic of a musician as reflected by their ceiling performance in discriminating all chordal conditions. Indeed, it may also be the case that the superiority of musicians over tone language speakers we observe for both behavioral tasks is linked to differences in auditory working

memory and attention which are typically heightened in musically trained individuals (Pallesen et al., 2010; Parbery-Clark, Skoe, Lam, et al., 2009; Strait et al., 2010).

#### 4.3. Dissociation between neural and perceptual representations of musical pitch

Our findings point to a dissociation between neural encoding of musical pitch and behavioral measures of pitch perception in tone language speakers (Fig. 5). Tone language experience may enhance brain mechanisms that subservise pitch processing in music as well as language. These mechanisms nevertheless may support distinct representations of pitch (Pfordresher & Brown, 2009, p. 1396). We demonstrate herein that enhanced brainstem representation of musical pitch notwithstanding (Fig. 3), it fails to translate into perceptual benefits for tone language speakers as it does for musicians (Figs. 4 and 5). It is plausible that dedicated brain circuitry exists to mediate musical percepts (e.g., Peretz, 2001) which may be differentially activated or highly developed dependent upon one's musical training (Foster & Zatorre, 2010; Pantev et al., 1998; Zarate & Zatorre, 2008). During pitch memory tasks, for example, non-musicians primarily recruit sensory cortices (primary/secondary auditory areas), while musicians show additional activation in the right supramarginal gyrus and superior parietal lobule, regions implicated in auditory short-term memory and pitch recall (Gaab & Schlaug, 2003). These results highlight the fact that similar neuronal networks are differentially recruited and tuned dependent on musical experience even under the same task demands.

Whether or not robust representations observed in subcortical responses engage such cortical circuitry may depend on the cognitive relevance of the stimulus to the listener (Abrams et al., 2010; Bidelman, Krishnan, et al., 2011; Chandrasekaran et al., 2009; Halpern, Martin, & Reed, 2008). Thus, in tone language speakers, it might be the case that information relayed from the brainstem fails to differentially engage these higher-level cortical mechanisms which subservise the perception and cognition of *musical pitch*, as it does for musicians. Indeed, we find no association between neurophysiological and behavioral measures in either group lacking musical training (C and NM) in contrast to the strong correlations between brain and behavior for musically trained listeners (Fig. 5). Our results, however, do not preclude the possibility that the brain-behavior dissociation observed in the Chinese group can be minimized or even eliminated with learning or training. That is, the neural enhancements we find in Chinese non-musicians, relative to English-speaking non-musicians (Fig. 3), could act to facilitate/accelerate improvements in their ability to discriminate musical pitch with active training.

As in language, brain networks engaged during music probably involve a series of computations applied to the neural representation at different stages of processing (Hickok & Poeppel, 2004; Poeppel, Idsardi, & van Wassenhove, 2008). Sensory information is continuously pruned and transformed along the auditory pathway dependent upon its functional role, e.g., whether or not it is linguistically- (Krishnan et al., 2009; Xu et al., 2006) or musically-relevant (Bidelman, Gandour, et al., 2011; Bidelman, Krishnan, et al., 2011) to the listener. Eventually, this information reaches complex cortical circuitry responsible for encoding/decoding musical percepts including melody and harmony (Koelsch & Jentschke, 2010) and the discrimination of pitch (Brattico et al., 2009; Koelsch, Schroger, & Tervaniemi, 1999; Tervaniemi, Just, Koelsch, Widmann, & Schroger, 2005). Yet, whether these encoded features are utilized effectively and retrieved during perceptual music tasks depends on the context and functional relevance of the sound, the degree of the listener's expertise, and the attentional demands of the task (Tervaniemi et al., 2005; Tervaniemi et al., 2009).

## 5. Conclusion

The presence of cross-domain transfer effects at pre-attentive stages of auditory processing does not necessarily imply that such effects will transfer to cognitive stages of processing invoked during music perception. Enhanced neural representation for pitch alone, while presumably necessary, is insufficient to produce perceptual benefits for music listening. We infer that neurophysiological enhancements are utilized by cognitive mechanisms only when the particular demands of an auditory environment coincide with one's long-term experience (Bidelman, Gandour, et al., 2011; Tervaniemi et al., 2009). Indeed, our data show that while both extensive music and language experience enhance neural representations for musical stimuli, only musicians truly make use of this information at a perceptual level.

## Acknowledgments

This article is based on part of a doctoral dissertation by the first author submitted to Purdue University in May 2011. Research supported in part by NIH R01 DC008549 (A.K.), NIDCD T32 DC00030 (G.B.), and Purdue University Bilsland dissertation fellowship (G.B.). Reprint requests should be addressed to Gavin Bidelman, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN, USA 47907-2038, or via email: gbidelma@purdue.edu.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bandc.2011.07.006.

## References

- Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2010). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*. doi:10.1093/cercor/bhq198.
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing and early reading ability in preschool children. *Journal of Experimental Child Psychology*, 83(2), 111–130.
- Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 97–103.
- Besson, M., Schon, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, 25(3–4), 399–410.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, 23(2), 425–434.
- Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, 29(42), 13165–13171.
- Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, 1355, 112–125.
- Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, 22(5), 212–216.
- Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European Journal of Neuroscience*, 33(3), 530–538.
- Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., & Purves, D. (2010). Major and minor music compared to excited and subdued speech. *Journal of the Acoustical Society of America*, 127(1), 491–503.
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B – Statistical Methodology*, 26(2), 211–252.
- Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Jarvenpaa, M., et al. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. *Journal of Cognitive Neuroscience*, 21(11), 2230–2244.
- Budge, H. (1943). *A study of chord frequencies*. New York: Teachers College, Columbia University.
- Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 215–264). San Diego: Academic Press.

- Burns, E. M., & Sampat, K. S. (1980). A note on possible culture-bound effects in frequency discrimination. *Journal of the Acoustical Society of America*, 68(6), 1886–1888.
- Burns, E. M., & Ward, W. D. (1978). Categorical perception – phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 63(2), 456–468.
- Carcagno, S., & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, 12(1), 89–100.
- Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature*, 396(6707), 128.
- Chandrasekaran, B., Gandour, J. T., & Krishnan, A. (2007). Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. *Restorative Neurology and Neuroscience*, 25(3–4), 195–210.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, 108(1), 1–9.
- Delogo, F., Lampis, G., & Olivetti Belardinelli, M. (2006). Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, 7(3), 203–207.
- Delogo, F., Lampis, G., & Olivetti Belardinelli, M. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, 22(1), 46–61.
- Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, 119(2), 719–722.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85(4), 341–354.
- Eberlein, R. (1994). *Die Entstehung der tonalen Klangsyntax [The origin of tonal-harmonic syntax]*. Frankfurt: Peter Lang.
- Foster, N. E., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, 20(6), 1350–1359.
- Franklin, M. S., Sledge Moore, K., Yip, C.-Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, 36(3), 353–365.
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, 16(6), 1010–1021.
- Gaab, N., & Schlaug, G. (2003). The effect of musicianship on pitch memory in performance matched groups. *Neuroreport*, 14(18), 2291–2295.
- Galbraith, G., Threadgill, M., Hemsley, J., Salour, K., Songdej, N., Ton, J., et al. (2000). Putative measure of peripheral and brainstem frequency-following in humans. *Neuroscience Letters*, 292, 123–127.
- Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149–175.
- Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12, 235–261.
- Gandour, J. T. (1994). Phonetics of tone. In R. Asher & J. Simpson (Eds.), *The encyclopedia of language & linguistics* (pp. 3116–3123). New York: Pergamon Press.
- Gao, E., & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21), 12663–12670.
- Halpern, A. R., Martin, J. S., & Reed, T. D. (2008). An ERP study of major–minor classification in melodies. *Music Perception*, 25(3), 181–191.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99.
- Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brain stem response to speech: A biological marker of auditory processing. *Ear and Hearing*, 26(5), 424–434.
- Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, 12(2), 125–143.
- Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical “language-network” serves the processing of music. *Neuroimage*, 17(2), 956–966.
- Koelsch, S., & Jentschke, S. (2010). Differences in electric brain responses to melodies and chords. *Journal of Cognitive Neuroscience*, 22(10), 2251–2262.
- Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309–1313.
- Krishnan, A. (2007). Human frequency following response. In R. F. Burkard, M. Don, & J. J. Eggermont (Eds.), *Auditory evoked potentials: Basic principles and clinical application* (pp. 313–335). Baltimore: Lippincott Williams & Wilkins.
- Krishnan, A., Bidelman, G. M., & Gandour, J. T. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, 268(1–2), 60–66.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, 110(3), 135–148.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, 23, 81–95.
- Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, 20(4), 408–413.
- Krohn, K. I., Brattico, E., Valimaki, V., & Tervaniemi, M. (2007). Neural representations of the hierarchical scale pitch structure. *Music Perception*, 24(3), 281–296.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America*, 124(5), 3235–3248.
- Lee, C. Y., & Lee, Y. F. (2010). Perception of musical pitch and lexical tones by Mandarin-speaking musicians. *Journal of the Acoustical Society of America*, 127(1), 481–490.
- Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences*, 9(1), 26–33.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2), 467–477.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: An MEG study. *Nature Neuroscience*, 4(5), 540–545.
- Marie, C., Kujala, T., & Besson, M. (2010a). Musical and linguistic expertise influence pre-attentive and attentive processing of non-speech sounds. *Cortex*. doi:10.1016/j.cortex.2010.1011.1006.
- Marie, C., Magne, C., & Besson, M. (2010b). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, 23(2), 294–305.
- Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, 19(9), 1453–1463.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219(1–2), 36–47.
- Moreno, S., & Besson, M. (2005). Influence of musical training on pitch processing: Event-related brain potential studies of adults and children. *Annals of the New York Academy of Sciences*, 1060, 93–97.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19(3), 712–723.
- Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, 241(1–2), 34–42.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., et al. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS One*, 5(6), e11120.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811–814.
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, 12(1), 169–174.
- 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(45), 14100–14107.
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30(6), 653–661.
- Patel, A. D. (2008). *Music, language, and the brain*. NY: Oxford University Press.
- Peretz, I. (2001). Brain specialization for music. New evidence from congenital amusia. *Annals of the New York Academy of Sciences*, 930, 153–165.
- Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, 71(6), 1385–1398.
- Plack, C. J., Oxenham, A. J., Fay, R. R., & Popper, A. N. (Eds.). (2005). *Pitch: Neural coding and perception* (Vol. 24). New York: Springer.
- Ping, L., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–210.
- Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 363(1493), 1071–1086.
- Schellenberg, E. G., & Peretz, I. (2008). Music, language and cognition: unresolved issues. *Trends in Cognitive Sciences*, 12(2), 45–46.
- Schellenberg, E. G., & Trehub, S. E. (2008). Is there an Asian advantage for pitch memory. *Music Perception*, 25(3), 241–252.
- Schon, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341–349.
- Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing*, 31(3), 302–324.
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, 17(8), 675–681.
- Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16(2), 374–381.
- Sohmer, H., Pratt, H., & Kinarti, R. (1977). Sources of frequency-following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology*, 42, 656–664.

- Stagray, J. R., & Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and nontone language. *Journal of Chinese Linguistics*, 21(1), 143–163.
- 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(1–2), 22–29.
- Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A*, 194(2), 169–183.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schroger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, 161(1), 1–10.
- Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, 30(8), 1636–1642.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, 4, 46–64.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422.
- Xu, Y. (2006). Tone in connected discourse. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 742–750). Oxford, UK: Elsevier.
- Xu, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Tong, Y., et al. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping*, 27(2), 173–183.
- Yip, M. (2002). *Tone*. New York: Cambridge University Press.
- Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage*, 40(4), 1871–1887.