COGNITIVE NEUROSCIENCE

Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch

Gavin M. Bidelman, Ananthanarayan Krishnan and Jackson T. Gandour
Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN, USA

Abstract

Important to Western tonal music is the relationship between pitches both within and between musical chords; melody and harmony are generated by combining pitches selected from the fixed hierarchical scales of music. It is of critical importance that musicians have the ability to detect and discriminate minute deviations in pitch in order to remain in tune with other members of their ensemble. Event-related potentials indicate that cortical mechanisms responsible for detecting mistuning and violations in pitch are more sensitive and accurate in musicians as compared with non-musicians. The aim of the present study was to address whether this superiority is also present at a subcortical stage of pitch processing. Brainstem frequency-following responses were recorded from musicians and non-musicians in response to tuned (i.e. major and minor) and detuned (± 4% difference in frequency) chordal arpeggios differing only in the pitch of their third. Results showed that musicians had faster neural synchronization and stronger brainstem encoding for defining characteristics of musical sequences regardless of whether they were in or out of tune. In contrast, non-musicians had relatively strong representation for major/minor chords but showed diminished responses for detuned chords. The close correspondence between the magnitude of brainstem responses and performance on two behavioral pitch discrimination tasks supports the idea that musicians' enhanced detection of chordal mistuning may be rooted at pre-attentive, sensory stages of processing. Findings suggest that perceptually salient aspects of musical pitch are not only represented at subcortical levels but that these representations are also enhanced by musical experience.

Introduction

Musical experience improves basic auditory acuity in both time and frequency as musicians are superior to non-musicians in perceiving and detecting rhythmic irregularities and fine-grained manipulations in pitch (Spiegel & Watson, 1984; Kishon-Rabin *et al.*, 2001; Micheyl *et al.*, 2006; Rammsayer & Altenmuller, 2006). Cortical event-related potentials offer neurophysiological evidence that musicians' perceptual advantages are probably due to sensory encoding enhancements of the pitch (Fujioka *et al.*, 2004; Tervaniemi *et al.*, 2009), timbre (Crummer *et al.*, 1994; Pantev *et al.*, 2001) and timing (Russeler *et al.*, 2001) of complex sounds. It is clear, then, that music-related functions rely heavily on cortical processing (e.g. Geiser *et al.*, 2009). Moreover, these reports also indicate that a musician's years of active engagement with complex auditory objects alter neurocognitive mechanisms and sharpen critical listening skills necessary for sophisticated music perception (for a review, see Tervaniemi, 2009).

Music performance necessitates the precise manipulation of pitch in order for an instrumentalist to remain in tune not only with him or herself, but also with surrounding members of the ensemble. As such,

it is critical that they detect deviations from the tempered scale in order to ensure proper musical tuning throughout a piece. As an index of cortical pitch discrimination, endogenous brain potentials (e.g. mismatched negativity) reveal that musicians automatically detect marginal pitch violations in musical sequences (e.g. detuned chords) otherwise undetectable for non-musicians (Koelsch *et al.*, 1999; Brattico *et al.*, 2002, 2009). However, whether this superior accuracy for pitch is exerted at pre-attentive levels in the cerebral cortex, or even at subcortical levels, is a matter of debate (cf. Tervaniemi *et al.*, 2005).

To index early stages of pre-attentive, subcortical pitch processing, we employed the scalp-recorded frequency-following response (FFR). The FFR reflects sustained phase-locked activity within the rostral brainstem, characterized by a periodic waveform that follows individual cycles of the stimulus (for review, see Krishnan, 2007). Use of the FFR has revealed that long-term music experience enhances brainstem representation of speech- (Wong *et al.*, 2007; Bidelman *et al.*, 2011; Bidelman & Krishnan, 2010) and musically-relevant (Musacchia *et al.*, 2007, 2008; Bidelman & Krishnan, 2009; Lee *et al.*, 2009) stimuli. When presented with a continuous pitch glide uncharacteristic of those found in music, Bidelman *et al.* (2011) found that musicians' FFRs showed selective enhancement for intermediate pitches of the diatonic musical scale. These findings demonstrate that musicians extract features of the auditory stream that help to define

Correspondence: Gavin M. Bidelman, as above.
E-mail: gbidelma@purdue.edu

melody and harmony, even at a subcortical level of processing (e.g. Tramo *et al.*, 2001; Bidelman & Krishnan, 2009).

Extending these results, we examine herein spectro-temporal properties of the FFR in response to tuned (i.e. major and minor) and detuned chordal arpeggios. Of specific interest is the effect on brainstem responses of parametrically manipulating the chordal third's pitch (in tune vs. out of tune). We predict musicians to show more robust brainstem representation for these defining features in musical pitch sequences providing a pre-attentive encoding scheme that may explain their superior pitch discrimination. In addition, it is hypothesized that both the neural encoding and perceptual performance will differ as a function of musical experience.

Materials and methods

Participants

Eleven English-speaking musicians (seven male, four female) and 11 non-musicians (six male, five female) were recruited from Purdue University to participate in the experiment. As determined by a music proficiency questionnaire, musically-trained participants (M) were amateur instrumentalists who had at least 10 years of continuous instruction on their principal instrument (mean ± SD; 12.4 ± 1.8 years), beginning at or before the age of 11 (8.7 ± 1.4 years). Each had formal private or group lessons within the past 5 years and currently played his∕her instrument (s). Non-musicians (NM) had no more than 1 year of formal music training (0.5 ± 0.5 years) on any combination of instruments in addition to not having received music instruction within the past 5 years (Table 1). All exhibited normal hearing sensitivity at octave frequencies between 500 and 4000 Hz and reported no previous history of neurological or psychiatric illnesses. There were no significant differences between the musician and non-musician groups in gender distribution ($P > 0.05$, Fisher's exact test). The two groups were also closely matched in age

(M: 22.63 ± 2.15 years, NM: 22.82 ± 3.40 years; $t_{20} = -0.15$, $P = 0.88$), years of formal education (M: 17.14 ± 1.76 years, NM: 16.55 ± 2.63 years; $t_{20} = 0.62$, $P = 0.54$) and handedness (laterality %, positive = right) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971) (M: 85.75 ± 15.9%, NM: 84.89 ± 20.99%; $t_{20} = 0.11$, $P = 0.91$). All participants were paid and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

Stimuli

Four triad arpeggios (i.e. three-note chords played sequentially) were constructed, which differed only in their chordal third (Fig. 1). Two sequences were exemplary arpeggios of Western music practice (major and minor chord); the other two represented detuned versions of these chords (detuned up and detuned down). Detuning was accomplished by manipulating the pitch of the chord's third such that it was either slightly sharp or flat of the actual major or minor third, respectively. 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. The fundamental frequencies (F0 s) of each of the three notes (i.e. root, third, fifth) per triad were as follows: major, 220, 277, 330 Hz; minor, 220, 262, 330 Hz; detuned up, 220, 287, 330 Hz;

TABLE 1. Musical background of participants

Participant	Instrument(s)	Years of training	Age of onset
Musicians			
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)
Non-musicians			
NM1	Piano	1	9
NM2	Clarinet	1	12
NM3	Piano	1	14
NM4	Flute	1	11
NM5	Guitar	0.5	15
NM6	Piano	1	10
NM7	–	0	–
NM8	–	0	–
NM9	–	0	–
NM10	–	0	–
NM11	–	0	–
Mean (SD)		0.50 (0.50)	11.8 (2.3)*

*Age-of-onset statistics for non-musicians were computed from the six participants with minimal musical training.
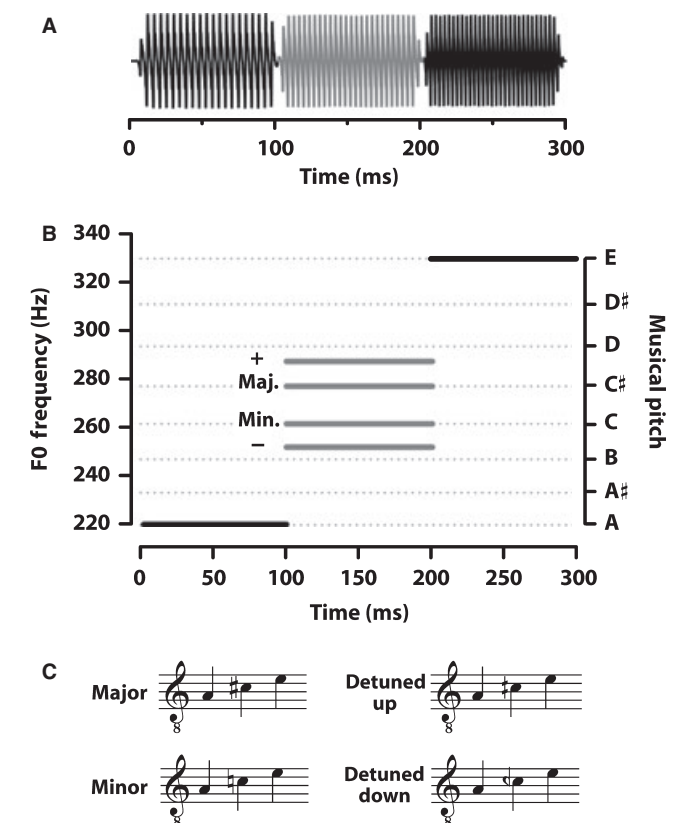

FIG. 1. Triad arpeggios used to evoke brainstem responses. (A) Four sequences were created by concatenating three 100 ms pitches together (B) whose F0 s corresponded to either prototypical (major, minor) or mistuned (detuned up, detuned down) versions of musical chords. Only the pitch of the chordal 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. (C) Musical notation for the four stimulus conditions.

and detuned down, 220, 252, 330 Hz. In the detuned arpeggios, mistuning in the chord's third represented a +4% or −4% difference in F0 from the actual major or minor third, respectively. A 4% deviation is greater than the just-noticeable difference for frequency (< 1%) (Moore, 2003) but smaller than a full musical semitone (6%). This amount of deviation is similar to previously published reports examining musicians' and non-musicians' cortical event-related potentials to detuned triads (e.g. Tervaniemi *et al.*, 2005; Brattico *et al.*, 2009). F0 s of the first and third notes (root and fifth) were identical across stimuli, i.e. 220 and 330 Hz, respectively. Thus, stimuli differed only in the pitch of their third (i.e. second note). Each note was 100 ms in duration including a 5 ms rise–fall time. For each sequence, the three notes were concatenated to create a contiguous chordal arpeggio of 300 ms duration. All stimuli were amplitude normalized to 80 dB sound pressure level.

Frequency-following response data acquisition

The brainstem is an essential relay along the auditory pathway that performs significant signal processing on sensory-level information before sending it on to the cortex. To assess early stages of subcortical auditory processing, we utilized the FFR, an evoked potential generated in the upper brainstem. Although it is possible that the far-field recorded FFR reflects concomitant activity of both cortical and subcortical structures, a number of studies have recognized the inferior colliculus (IC) of the brainstem as its primary neural generator. This arises from the fact that (i) the shorter latency of the FFR ($\sim$7–12 ms) activity is too early to reflect a contribution from cortical generators (Galbraith *et al.*, 2000), (ii) there is a high correspondence between far-field FFR and near-field intra-cranial potentials recorded directly from the IC (Smith *et al.*, 1975), (iii) the FFR is abolished following cryogenic cooling of the IC (Smith *et al.*, 1975) and, lastly, (iv) the FFR is absent with brainstem lesions confined to the IC (Sohmer & Pratt, 1977).

The FFR recording protocol was similar to that used in previous reports from our laboratory (Bidelman & Krishnan, 2009; Krishnan *et al.*, 2009). 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) and to ignore the sound that they heard. Subjects were allowed to sleep throughout the duration of the FFR experiment ($\sim$80% fell asleep). FFRs were recorded from each participant in response to monaural stimulation of the right ear at an intensity of 80 dB sound pressure level 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. The 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) using a sampling rate of 10 kHz.

The continuous electroencephalogram was recorded differentially between Ag–AgCl scalp electrodes placed on the midline of the forehead at the hairline (non-inverting, active) and right mastoid (A2; inverting, reference). Another electrode placed on the mid-forehead served as the common ground. Such a vertical electrode montage provides the optimal configuration for recording brainstem responses (Galbraith *et al.*, 2000). Inter-electrode impedances were maintained at $\leq$ 1 kΩ, amplified by 200 000 and filtered online between 30 and 5000 Hz. A total of 3000 artifact-free sweeps were recorded for each run lasting approximately 20 min. The electroencephalograms were stored to hard disk for offline processing. Raw electroencephalograms were then divided into epochs using an analysis time window from 0 to 320 ms (0 ms is stimulus onset). FFRs were extracted by time-domain averaging each epoch over the duration of the recording. Sweeps containing activity exceeding ± 35 μV were rejected as artifacts and excluded from the final average. FFR response waveforms were further band-pass filtered from 100 to 2500 Hz (−6 dB/octave roll-off) to minimize low-frequency physiologic noise and limit the inclusion of cortical activity. In total, each FFR response waveform represents the average of 3000 artifact-free trials over a 320 ms acquisition window.

Frequency-following response data analysis

Neural latencies to note onsets

To quantify the temporal precision of each response, onset latencies were measured within the FFR corresponding to each note of the major chord stimulus. The onset of sustained phase-locking in the FFR can be represented by a large negative deflection occurring between 15 and 20 ms post-stimulus onset (e.g. Musacchia *et al.*, 2007; Strait *et al.*, 2009). As such, the latency of the largest negative trough in this time window was taken as the onset of neural activity in response to the chord sequence (i.e. the onset of the first note). The latency of the positive peak immediately preceding this negative marker was also measured. Subsequent note onsets were recorded using identical criteria in the expected time windows predicted from the length of notes (100 ms) in the stimulus, i.e. note 2, $\sim$115–120 ms and note 3, $\sim$215–220 ms. Peaks were identified by G.M.B. and confirmed by another observer experienced in electrophysiology who was blind to the participant's group. Inter- and intra-observer reliabilities for onset latency selections were > 97%. The difference between the positive–negative onset peak latencies (i.e. P-N onset duration, expressed in ms) was taken as an index of the neural synchronization to each musical note. A longer P-N onset duration indicates slower, more sluggish neural synchronization to each note in the auditory stream, whereas shorter durations indicate more precise, time-locked neural activity to each note of the stimulus.

Brainstem response fundamental frequency magnitudes

The FFR pitch encoding magnitude was quantified by measuring the F0 component from each response waveform for each of the three notes per melodic triad. FFRs were segmented into three 100 ms sections (15–115, 115–215 and 215–315 ms) corresponding to the sustained portions of the response to each musical note. The spectrum of each response segment was computed by taking the Fast Fourier Transform (FFT) of a time-windowed version of its temporal waveform (Gaussian window, 1 Hz resolution). For each subject per arpeggio and note, the magnitude of F0 was measured as the peak in the FFT, relative to the noise floor, which fell in the same frequency range as the F0 of the input stimulus (note 1: 210–230 Hz; note 2: 245–300 Hz; note 3: 320–340 Hz; see stimulus F0 tracks, Fig. 1B). All FFR data analyses were performed using custom routines coded in MATLAB® 7.9 (The MathWorks, Inc., Natick, MA, USA).

Behavioral measure of chordal detuning

A pitch discrimination task was conducted to determine whether musicians and non-musicians differed in their ability to detect chordal detuning at a perceptual level. Five musicians and five non-musicians who also took part in the FFR experiment participated in the

behavioral task. Discrimination sensitivity was measured separately for the three most meaningful stimulus pairings (major/detuned up, minor/detuned down, major/minor) using a same–different task. For each of these three conditions, participants heard 100 pairs of the chordal arpeggios presented with an inter-stimulus 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 for each participant 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 an intensity of $\sim$75 dB sound pressure level through circumaural headphones (HD 580; Sennheiser Electronic Corp., Old Lyme, CT, USA). Stimulus presentation and response collection were implemented in a custom graphical user interface coded in MATLAB.

Statistical analysis

A two-way, mixed-model ANOVA (SAS®; SAS Institute, Inc., Cary, NC, USA) was conducted on F0 magnitudes derived from FFRs in order to evaluate the effects of musical experience and context (i.e. prototypical vs. non-prototypical sequence) on brainstem encoding of musical pitch. Group (two levels: musicians, non-musicians) functioned as the between-subjects factor and stimulus (four levels: major, minor, detuned up, detuned down) as the within-subjects factor. The magnitudes of F0 encoding for the first and last note in the stimuli (i.e. chord root and fifth) were not analyzed statistically given that, by design, these components did not differ in the input stimuli themselves and, moreover, responses showed no observable differences between stimuli (see Fig. S1).

The duration of the FFR P-N onset complex was analyzed using a similar model with group (two levels: musicians, non-musicians) as the between-subjects factor and notes (three levels: first, second, third) as the within-subjects factor.

Behavioral discrimination sensitivity scores (d') were computed using hit (H) and false alarm (FA) rates [i.e. $d' = z(H) - z(FA)$, where $z(.)$ represents the z-score operator]. Two musicians obtained perfect accuracy (FA = 0) implying a d' of infinity. In these cases, a correction was applied by adding 0.5 to both the number of hits and false alarms in order to compute a finite d' (Macmillan & Creelman, 2005). Based on initial diagnostics and the Box-Cox procedure (Box & Cox, 1964), d' scores were log-transformed to improve the normality and homogeneity of variance assumptions necessary for a parametric ANOVA. Log-transformed d' scores were submitted to a two-way mixed model with group (two levels: musicians, non-musicians) as the between-subjects factor and stimulus pair (three levels: major/detuned up, minor/detuned down, major/minor) 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 ($\alpha_{individual} = 0.0167$). Where appropriate, partial eta-squared ($\eta^2_{partial}$) values are reported to indicate effect sizes.

Results

Neural latencies to note onsets

Visual inspection indicated that, within each group, there were no latency differences between arpeggios. Thus, only results for the major chord are presented here. Grand average FFR time-waveforms in

response to the major chord stimulus are shown per group in Fig. 2A. For both groups, clear onset components are seen at the three time marks corresponding to the individual onset of each note (i.e. large negative deflections; note $1 \approx 17$ ms, note $2 \approx 117$ ms, note $3 \approx 217$ ms). Relative to non-musicians, musicians showed larger amplitudes across the duration of their response. This amplified neural activity was most evident throughout the chordal third (i.e. second note, $\sim$110–210 ms), the defining pitch of the sequence. Within this same time window, non-musicians' responses showed a reduced amplitude, indicating poorer representation of this chord-defining pitch (see also Fig. 3).

Neural onset synchrony, as measured by the duration of the P-N onset complex, was observed to be more robust with earlier onset response components for musicians (Fig. 2B and C). An omnibus ANOVA on P-N onset duration revealed significant main effects of group ($F_{1,20} = 14.17$, $P = 0.0012$, $\eta^2_{partial} = 0.41$) and note ($F_{2,40} = 5.01$, $P = 0.0114$, $\eta^2_{partial} = 0.20$). By group, post-hoc Bonferroni-adjusted multiple comparisons revealed that the P-N onset duration was identical across notes for musicians ($P > 0.05$) but that it increased from the first to last note for non-musicians ($P = 0.01$) (Fig. 2C). The widening of the P-N onset complex with each subsequent note can be attributed to the increased prolongation (i.e. larger absolute latency) of each negative peak relative to the positive portion of the onset response (% increase from note 1 to 3: $M_{pos} = 2.54\%$, $NM_{pos} = 2.98\%$, $M_{neg} = 3.03\%$, $NM_{neg} = 4.24\%$). Compared with non-musicians, the relatively shorter duration of musicians' onsets across notes indicates more precise, time-locked neural activity to each musical note.

Brainstem response fundamental frequency magnitudes of chordal thirds

The FFR encoding of F0 for the thirds of chordal standard and detuned arpeggios are shown in Fig. 3. Individual panels show the meaningful comparisons that fall within the range of a semitone: A, major vs. minor; B, major vs. detuned up; C, minor vs. detuned down. An omnibus ANOVA on F0 encoding revealed significant main effects of group ($F_{1,20} = 33.31$, $P < 0.001$, $\eta^2_{partial} = 0.62$) and stimulus ($F_{3,60} = 8.00$, $P = 0.0001$, $\eta^2_{partial} = 0.29$) on F0 encoding, as well as a group $\times$ stimulus interaction ($F_{3,60} = 3.11$, $P = 0.0331$, $\eta^2_{partial} = 0.13$).

A priori contrasts revealed that, regardless of the arpeggio, musicians' brainstem responses contained a larger F0 magnitude than those of non-musicians ($P \leq 0.01$) (Fig. 3A–C). By group, the F0 magnitude did not differ across triads for musicians ($P > 0.05$), indicating superior encoding regardless of whether the chordal third was major or minor, in or out of tune. Interestingly, for non-musicians, F0 encoding was identical between the major and minor chords ($P > 0.05$) (Fig. 3A), two of the most regularly occurring sequences in music (Budge, 1943), but was significantly reduced for the detuned sequences ($P < 0.01$) (Fig. 3B and C). Together, these results indicate superior encoding of pitch-relevant information in musicians regardless of chordal temperament, and that brainstem encoding is disrupted with chordal detuning only in the non-musician group.

Behavioral chordal third discrimination

Group behavioral discrimination sensitivity scores, as measured by d', are shown for musicians and non-musicians in Fig. 4. Values represent the 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 $d' = 0$ represents chance
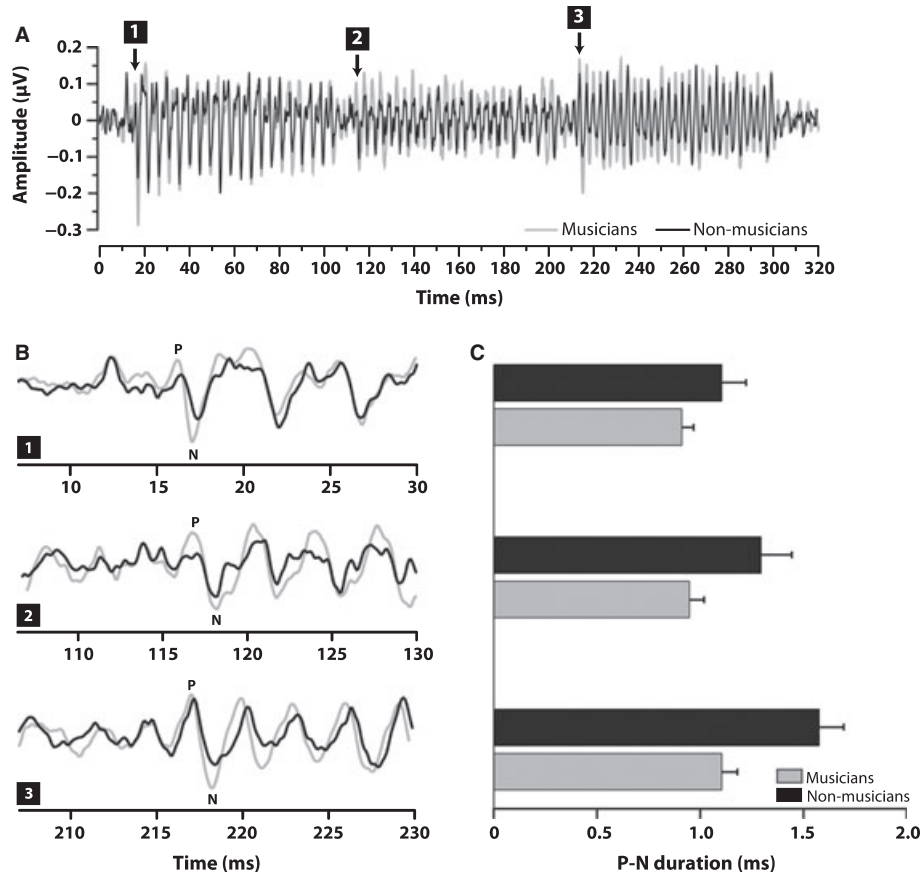
FIG. 2. FFR onset latencies to notes in the major chord stimulus. (A) Grand average FFR time-waveforms per group. Relative to non-musicians, musicians show larger amplitudes across the duration of their responses but most especially throughout the chordal third (i.e. second note), the defining pitch of the sequence. Neural onsets to individual notes are demarcated by their respective number (1–3). (B) Expanded time windows around onset responses to individual notes (note 1 ≈ 17 ms, note 2 ≈ 117 ms, note 3 ≈ 217 ms). Relative to non-musicians, musicians generally show larger peak amplitudes in their P-N onset complexes. (C). In addition, the shorter durations of musicians' P-N onset complex across notes indicate their more precise, time-locked neural activity to each musical pitch. Error bars, 1 SE; P-N, difference between positive–negative onset peak latencies.

performance. An ANOVA on d' scores revealed significant main effects of group ($F_{1,8} = 31.70$, $P = 0.0005$, $\eta^2_{\mathrm{partial}} = 0.80$) and stimulus pair ($F_{2,16} = 5.93$, $P = 0.0118$, $\eta^2_{\mathrm{partial}} = 0.43$), as well as a group × stimulus pair interaction ($F_{2,16} = 4.48$, $P = 0.0284$, $\eta^2_{\mathrm{partial}} = 0.36$). Multiple comparisons revealed that musicians performed equally well above threshold for all conditions and did not differ in their discrimination ability between standard (major/minor) and detuned (major/up, minor/down) stimulus pairings. In contrast, non-musicians only obtained suprathreshold performance when discriminating the major/minor pair and could not accurately distinguish detuned chords from the major or minor standards. These results indicate that musicians perceive minute changes in musical pitch that are otherwise undetectable by non-musicians (see also, F0 difference limens in Fig. S2).

Discussion

There are two major findings of this study. First, compared with non-musicians, musicians had faster neural synchronization and stronger brainstem encoding for the third of triadic arpeggios (the defining feature of the chord) regardless of whether the sequence was in or out of tune. Non-musicians, however, had stronger encoding for the prototypical major and minor chords than detuned chords. Second, musicians showed a superior performance over non-musicians in discriminating standard and detuned arpeggios as well as simple pitch

change detection (i.e. F0 difference limens), indicating that extensive musical training sharpens perceptual mechanisms operating on pitch. Close correspondence between the pattern of brainstem response magnitudes and performance in perceptual pitch discrimination tasks supports the idea that musicians' enhanced detection of chordal detuning may be rooted at pre-attentive, sensory stages of processing.

Neural basis for musicians' enhancements: a product of subcortical plasticity

Our findings provide further evidence for experience-dependent plasticity induced by long-term music experience (Tervaniemi *et al.*, 1997; Munte *et al.*, 2002; Zatorre & McGill, 2005; Tervaniemi, 2009; Kraus & Chandrasekaran, 2010). Across all stimuli, musicians had faster neural synchronization (Fig. 2) and stronger brainstem encoding (Figs 3 and S1) for the third of triadic arpeggios, the defining feature of the chord. From a neurophysiologic perspective, the optimal encoding that we find in musicians reflects enhancement in phase-locked activity within the rostral brainstem. IC architecture (Schreiner & Langner, 1997; Braun, 1999) and its response properties (Langner, 1981, 1997) provide optimal hardware in the midbrain for extracting complex pitch. Such mechanisms are especially well suited for the encoding of pitch relationships recognized by music (e.g. major/minor chords) over those which are less harmonic and, consequently, out of tune (e.g. detuned chords) (Braun, 2000; Lots & Stone, 2008).
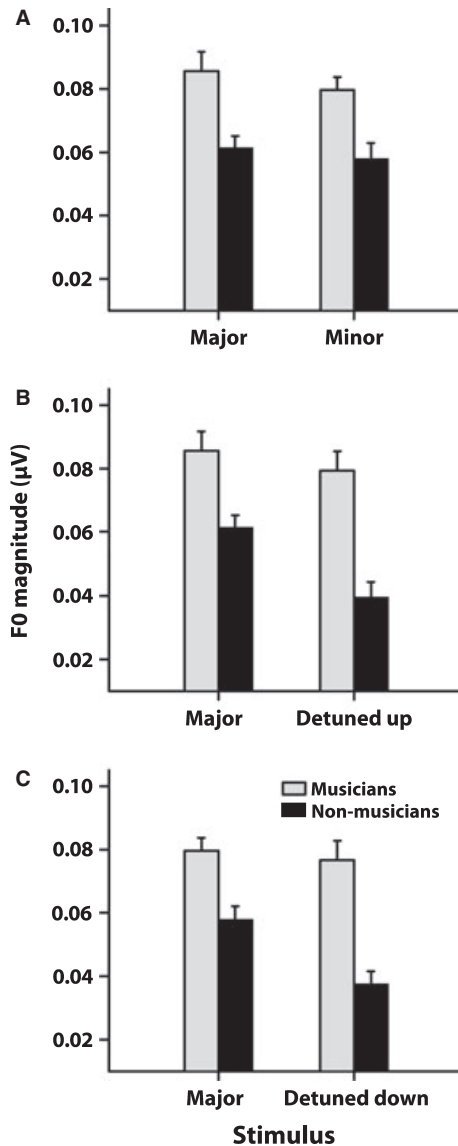
FIG. 4. Behavioral group d' scores for discriminating chord arpeggios. By convention, discrimination threshold is represented by $d' = 1$ (dashed line). Musicians show superior performance (well above threshold) in discriminating all chord pairings, including standard chords in music from versions in which the third is out of tune (i.e. major/up and minor/down). Non-musicians, however, only discriminate the major/minor pairing above threshold and are unable to accurately differentiate standard musical chords from their detuned versions (i.e. subthreshold discrimination, major/up and minor/down).

FIG. 3. Musicians show enhanced pitch encoding of chordal thirds in response to both prototypical (A) and detuned (B and C) arpeggios. Pitch encoding is defined as the spectral magnitude of F0 measured from FFR responses. (A) Within both groups, no differences are seen between F0 magnitudes for the major and minor third, probably due to their overabundance in Western music. However, musicians show enhanced representation for these defining musical notes relative to their non-musician counterparts. When the third of the chord is slightly sharp (+4%) or flat (−4%) relative to the major and minor third, musicians show invariance in their encoding, representing detuned notes as well as the tempered pitches (B and C). In contrast, non-musicians show marked decrease in representation of F0 when the chord is detuned from the standard major or minor prototype.

The enhancements in musicians represent a strengthening of this subcortical circuitry developed from many hours of active exposure to the dynamic spectro-temporal properties found in music.

Although the FFR primarily reflects an aggregate of neural activity generated in the midbrain (Smith *et al.*, 1975; Sohmer & Pratt, 1977; Galbraith *et al.*, 2000; Akhoun *et al.*, 2008), this does not preclude the possibility that the superiority that we observe in musicians may reflect activity already enhanced by lower level structures (i.e. cochlea or caudal brainstem nuclei). Studies examining otoacoustic emissions 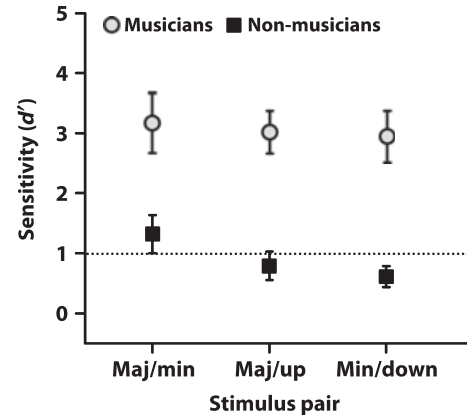have consistently shown larger contralateral suppression effects in musicians, suggesting that musical training strengthens medial efferent feedback from the caudal brainstem (superior olivary complex) to the cochlea (Micheyl *et al.*, 1997; Perrot *et al.*, 1999; Brashears *et al.*, 2003). Given the putative connection between this cochlear active process and behavioral pitch discrimination sensitivity (Norena *et al.*, 2002), it is conceivable that the behavioral and physiological superiority that we find in musicians may result from enhancements beginning even as early as the cochlea.

Musicians' brainstem responses are less susceptible to detuning than those of non-musicians

Of particular importance to Western tonal music is the relationship between pitches both within and between musical chords. Melody and harmony are generated by combining pitch combinations selected from the fixed hierarchical scales of music. Indeed, typified by our stimuli (Fig. 1), single pitches can determine the quality (e.g. major vs. minor) and temperament (i.e. in vs. out of tune) of musical pitch sequences. As such, ensemble performance requires that musicians constantly monitor pitch in order to produce correct musical quality and temperament relative to themselves, as well as with the entire ensemble. The ability to detect and discriminate minute deviations in pitch is therefore of critical importance to both the performance and appreciation of tonal music.

Across all chordal arpeggios, musicians showed enhanced pitch encoding over non-musician controls, suggesting that extensive music experience magnifies sensory-level representation of musically relevant stimuli (for enhancements to speech-relevant stimuli, see Wong *et al.*, 2007; Bidelman *et al.*, 2011; Parbery-Clark *et al.*, 2009; Strait *et al.*, 2009; Bidelman & Krishnan, 2010). Major or minor, in or out of tune, we found that musicians' FFRs showed no appreciable reduction in neural representation of pitch with parametric manipulation of the chordal third (Fig. 3B and C). An encoding scheme of this nature (which represents both in and out of tune pitch equally) would be extremely advantageous for a musician. Phase-locked activity generated in the brainstem is eventually relayed to cortical mechanisms responsible for detecting and discriminating violations in pitch. Feeding this type of circuitry with stronger subcortical information

would provide such mechanisms with a more robust representation for pitch regardless of its tuning characteristics. Indeed, musicians show enhancements in the earliest stages of cortical processing, suggesting that more robust sensory information is input to the auditory cortex (Baumann *et al.*, 2008). Stronger representations throughout the auditory pathway would, in turn, enable pitch change detection mechanisms (e.g. mismatched negativity generators) to operate more efficiently and accurately in musicians. Transformations of this sort may underlie the enhancements observed in musicians' cortical responses to violations in musical pitch (including chords), which are otherwise undetectable for non-musicians (Koelsch *et al.*, 1999; Brattico *et al.*, 2002, 2009; Schon *et al.*, 2004; Moreno & Besson, 2005; Magne *et al.*, 2006; Nikjeh *et al.*, 2008).

In contrast to musicians, brainstem responses of non-musicians were differentially affected by the musical context of arpeggios (i.e. in vs. out of tune), resulting in diminished magnitudes for detuned chords relative to their major/minor counterparts (Fig. 3). Although the source of such differential group effects is not entirely clear, both neurophysiological and experience-driven mechanisms may account for our observations. Differences in loudness adaptation, probably mediated by caudal brainstem efferents, have been reported between groups, suggesting that a musician's auditory system maintains the intensity of sound more faithfully over time than in non-musicians (Micheyl *et al.*, 1995). A reduction in adaptation, for example, may partially explain the invariance of musicians' FFR amplitude across musical notes (Fig. 2A) and their more efficient neural synchronization as compared with the weaker, more sluggish responses of non-musicians (Fig. 2B and C). Physiologic explanations notwithstanding, the more favorable encoding of prototypical musical sequences may be likened to the fact that even non-musicians are experienced listeners with certain chords (e.g. Bowling *et al.*, 2010). Major and minor triads are among the most commonly occurring chords in tonal music (Budge, 1943; Eberlein, 1994). Over the course of a lifetime, exposure to the stylistic norms of Western music may tune brain mechanisms to respond to the more probable pitch relationships found in music (e.g. Loui *et al.*, 2009, 2010). Indeed, we find that chords that do not intentionally occur in music practice (e.g. our detuned chords) elicit weaker responses from non-musician participants (compare Fig. 3A with Fig. 3B and C). These results are consistent with the notion that pre-attentive pitch-change processing is generally enhanced, even for non-musicians, in familiar musical contexts (e.g. major/minor) (Koelsch *et al.*, 2000; Brattico *et al.*, 2002). In addition, these data converge with the observation that, at the level of the brainstem, musically dissonant pitch relationships (e.g. detuned chords) elicit weaker neural responses than consonant relationships (e.g. major/minor chords) for musically untrained individuals (Bidelman & Krishnan, 2009).

Brain–behavior relationship for pitch discrimination

We found that musicians, relative to non-musicians, were superior at detecting subtle changes in fixed pitch (Fig. S2) as well as discriminating detuned arpeggios from standards (Fig. 4). These psychophysical data corroborate previous reports showing that long-term musical training heightens behavioral sensitivity to subtle nuances in pitch (Spiegel & Watson, 1984; Pitt, 1994; Kishon-Rabin *et al.*, 2001; Tervaniemi *et al.*, 2005; Micheyl *et al.*, 2006; Nikjeh *et al.*, 2008; Strait *et al.*, 2010). Of particular interest here is the fact that only musicians were able to discriminate standard and detuned arpeggios above threshold (Fig. 4). Non-musicians, however, could only reliably discriminate major from minor chords; their performance in

distinguishing detuned chords from standards fell below threshold. These results indicate that musicians perceive fine-grained changes in musical pitch, both in isolated static notes and time-varying sequences, which are otherwise undetectable for non-musicians.

Parallel results were seen in brainstem responses. We found that musicians' FFR pitch encoding was impervious to changes in the tuning characteristics of the eliciting arpeggio (Fig. 3) and, correspondingly, they reached ceiling performance in arpeggio discrimination across all conditions (Fig. 4). Non-musicians, who showed poorer encoding for detuned relative to standard arpeggios, subsequently were unable to detect chordal detuning. The close correspondence between brainstem F0 magnitude and behavioral performance suggests that musicians' enhanced detection of chordal detuning may be rooted in pre-attentive, sensory stages of processing. Indeed, it is suggested that, in musicians, perceptual decision mechanisms related to pitch may use pre-attentively encoded neural information more efficiently than in non-musicians (Tervaniemi *et al.*, 2005).

To date, only a few studies have investigated the role of subcortical processing in forming the perceptual attributes related to musical pitch (Tramo *et al.*, 2001; Bidelman & Krishnan, 2009; Lee *et al.*, 2009). Enhancements in cortical processing can account for musicians' improved perceptual discrimination of pitch (Koelsch *et al.*, 1999; Tervaniemi *et al.*, 2005; Brattico *et al.*, 2009). However, the extant literature is unclear whether this superiority depends on attention (Tervaniemi *et al.*, 2005; Halpern *et al.*, 2008) or also manifests at a pre-attentive level (Tervaniemi *et al.*, 1997; Koelsch *et al.*, 1999). Utilizing the pre-attentive brainstem FFR, our results suggest that this superior ability may emerge well before cortical involvement. As in language (Hickok & Poeppel, 2004), brain networks engaged during music probably involve a series of computations applied to the neural representation at different stages of processing (e.g. Bidelman *et al.*, 2011). Physical acoustic periodicity is transformed to musically relevant neural periodicity very early along the auditory pathway (auditory nerve) (Tramo *et al.*, 2001), and transmitted and enhanced in subsequently higher levels in the auditory brainstem (Bidelman & Krishnan, 2009; present study). Eventually, this information reaches the complex cortical architecture responsible for generating and controlling musical percepts including melody/harmony (Koelsch & Jentschke, 2010) and the discrimination of pitch (Koelsch *et al.*, 1999; Tervaniemi *et al.*, 2005; Brattico *et al.*, 2009). We argue that abstract representations of musical pitch are grounded in sensory features that emerge very early along the auditory pathway.

Conclusions

Our findings demonstrate that musicians, relative to non-musicians, have faster onset neural synchronization and stronger encoding of defining characteristics of musical pitch sequences in the auditory brainstem. These results show that the auditory brainstem is not hard-wired, but rather is changed by an individual's training and/or listening experience. The close correspondence between brainstem responses and discrimination performance supports the idea that enhanced representation of perceptually salient aspects of musical pitch may be rooted subcortically at a sensory stage of processing. Traditionally neglected in discussions of the neurobiology of music, we find that the brainstem plays an active role in not only the neural encoding of musically relevant sound but probably influences later processes governing music perception. Our findings further show that musical expertise modulates pitch encoding mechanisms that are not under direct attentional control (cf. Tervaniemi *et al.*, 2005).

Supporting Information

Additional supporting information may be found in the online version of this article:

Fig. S1. FFR pitch encoding magnitude for each note of the four chordal arpeggio stimuli: major, minor, detuned up, detuned down.

Fig. S2. Behavioral frequency difference limens (F0 DLs) for musicians and non-musicians.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset by Wiley-Blackwell. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Acknowledgements

Abbreviations

F0, fundamental frequency; FFR, frequency-following response; FFT, Fast Fourier Transform; IC, inferior colliculus; M, musicians; NM, non-musicians; P-N, positive–negative.

References

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

Baumann, S., Meyer, M. & Jancke, L. (2008) Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *J. Cogn. Neurosci.*, **20**, 2238–2249.

Bidelman, G.M. & Krishnan, A. (2009) Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.*, **29**, 13165–13171.

Bidelman, G.M. & Krishnan, A. (2010) Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.*, **1355**, 112–125.

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. *J. Cogn. Neurosci.*, **23**, 425–434.

Bowling, D.L., Gill, K., Choi, J.D., Prinz, J. & Purves, D. (2010) Major and minor music compared to excited and subdued speech. *J. Acoust. Soc. Am.*, **127**, 491–503.

Box, G.E.P. & Cox, D.R. (1964) An analysis of transformations. *J. Roy. Stat. Soc. B*, **26**, 211–252.

Brashears, S.M., Morlet, T.G., Berlin, C.I. & Hood, L.J. (2003) Olivocochlear efferent suppression in classical musicians. *J. Am. Acad. Audiol.*, **14**, 314–324.

Brattico, E., Naatanen, R. & Tervaniemi, M. (2002) Context effects on pitch perception in musicians and nonmusicians: evidence from event-related-potential recordings. *Music Percept.*, **19**, 199–222.

Brattico, E., Pallesen, K.J., Varyagina, O., Bailey, C., Anourova, I., Jarvenpaa, M., Eerola, T. & Tervaniemi, M. (2009) Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. *J. Cogn. Neurosci.*, **21**, 2230–2244.

Braun, M. (1999) Auditory midbrain laminar structure appears adapted to f0 extraction: further evidence and implications of the double critical bandwidth. *Hear. Res.*, **129**, 71–82.

Braun, M. (2000) Inferior colliculus as candidate for pitch extraction: multiple support from statistics of bilateral spontaneous otoacoustic emissions. *Hear. Res.*, **145**, 130–140.

Budge, H. (1943) *A Study of Chord Frequencies*. Teachers College, Columbia University, New York, NY.

Crummer, G.C., Walton, J.P., Wayman, J.W., Hantz, E.C. & Frisina, R.D. (1994) Neural processing of musical timbre by musicians, nonmusicians, and musicians possessing absolute pitch. *J. Acoust. Soc. Am.*, **95**, 2720–2727.

Eberlein, R. (1994) *Die Entstehung der tonalen Klangsyntax*. Peter Lang, Frankfurt.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R. & Pantev, C. (2004) Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.*, **16**, 1010–1021.

Galbraith, G., Threadgill, M., Hemsley, J., Salour, K., Songdej, N., Ton, J. & Cheung, L. (2000) Putative measure of peripheral and brainstem frequency-following in humans. *Neurosci. Lett.*, **292**, 123–127.

Geiser, E., Ziegler, E., Jancke, L. & Meyer, M. (2009) Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, **45**, 93–102.

Halpern, A.R., Martin, J.S. & Reed, T.D. (2008) An ERP study of major-minor classification in melodies. *Music Percept.*, **25**, 181–191.

Hickok, G. & Poeppel, D. (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, **92**, 67–99.

Kishon-Rabin, L., Amir, O., Vexler, Y. & Zaltz, Y. (2001) Pitch discrimination: are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.*, **12**, 125–143.

Koelsch, S. & Jentschke, S. (2010) Differences in electric brain responses to melodies and chords. *J. Cogn. Neurosci.*, **22**, 2251–2262.

Koelsch, S., Schroger, E. & Tervaniemi, M. (1999) Superior pre-attentive auditory processing in musicians. *Neuroreport*, **10**, 1309–1313.

Koelsch, S., Gunter, T., Friederici, A.D. & Schroger, E. (2000) Brain indices of music processing: "nonmusicians" are musical. *J. Cogn. Neurosci.*, **12**, 520–541.

Kraus, N. & Chandrasekaran, B. (2010) Music training for the development of auditory skills. *Nat. Rev. Neurosci.*, **11**, 599–605.

Krishnan, A. (2007) Human frequency following response. In Burkard, R.F., Don, M. & Eggermont, J.J. (Eds), *Auditory Evoked Potentials: Basic Principles and Clinical Application*. Lippincott Williams & Wilkins, Baltimore, MD, pp. 313–335.

Krishnan, A., Gandour, J.T., Bidelman, G.M. & Swaminathan, J. (2009) Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, **20**, 408–413.

Langner, G. (1981) Neuronal mechanisms for pitch analysis in the time domain. *Exp. Brain Res.*, **44**, 450–454.

Langner, G. (1997) Neural processing and representation of periodicity pitch. *Acta Otolaryngol. Suppl.*, **532**, 68–76.

Lee, K.M., Skoe, E., Kraus, N. & Ashley, R. (2009) Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.*, **29**, 5832–5840.

Lots, I.S. & Stone, L. (2008) Perception of musical consonance and dissonance: an outcome of neural synchronization. *J. R. Soc. Interface*, **5**, 1429–1434.

Loui, P., Wu, E.H., Wessel, D.L. & Knight, R.T. (2009) A generalized mechanism for perception of pitch patterns. *J. Neurosci.*, **29**, 454–459.

Loui, P., Wessel, D.L. & Hudson Kam, C.L. (2010) Humans rapidly learn grammatical structure in a new musical scale. *Music Percept.*, **27**, 377–388.

Macmillan, N.A. & Creelman, C.D. (2005) *Detection Theory: A User's Guide*. Lawrence Erlbaum Associates, Inc., Mahwah, N.J.

Magne, C., Schon, D. & Besson, M. (2006) Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J. Cogn. Neurosci.*, **18**, 199–211.

Micheyl, C., Carbonnel, O. & Collet, L. (1995) Medial olivocochlear system and loudness adaptation: differences between musicians and non-musicians. *Brain Cogn.*, **29**, 127–136.

Micheyl, C., Khalfa, S., Perrot, X. & Collet, L. (1997) Difference in cochlear efferent activity between musicians and non-musicians. *Neuroreport*, **8**, 1047–1050.

Micheyl, C., Delhommeau, K., Perrot, X. & Oxenham, A.J. (2006) Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.*, **219**, 36–47.

Moore, B.C.J. (2003) *An Introduction to the Psychology of Hearing*. Academic Press, Amsterdam, Boston.

Moreno, S. & Besson, M. (2005) Influence of musical training on pitch processing: event-related brain potential studies of adults and children. *Ann. NY Acad. Sci.*, **1060**, 93–97.

Munte, T.F., Altenmuller, E. & Jancke, L. (2002) The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.*, **3**, 473–478.

Musacchia, G., Sams, M., Skoe, E. & Kraus, N. (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl Acad. Sci. USA*, **104**, 15894–15898.

Musacchia, G., Strait, D. & Kraus, N. (2008) Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.*, **241**, 34–42.

Nikjeh, D.A., Lister, J.J. & Frisch, S.A. (2008) Hearing of note: an electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology*, **45**, 994–1007.

Norena, A., Micheyl, C., Durrant, J., Chery-Croze, S. & Collet, L. (2002) Perceptual correlates of neural plasticity related to spontaneous otoacoustic emissions? *Hear. Res.*, **171**, 66–71.

Oldfield, R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.

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. *J. Neurosci.*, **29**, 14100–14107.

Perrot, X., Micheyl, C., Khalfa, S. & Collet, L. (1999) Stronger bilateral efferent influences on cochlear biomechanical activity in musicians than in non-musicians. *Neurosci. Lett.*, **262**, 167–170.

Pitt, M.A. (1994) Perception of pitch and timbre by musically trained and untrained listeners. *J. Exp. Psychol. Hum. Percept. Perform.*, **20**, 976–986.

Rammsayer, T. & Altenmuller, E. (2006) Temporal information processing in musicians and nonmusicians. *Music Percept.*, **24**, 37–48.

Russeler, J., Altenmuller, E., Nager, W., Kohlmetz, C. & Munte, T.F. (2001) Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neurosci. Lett.*, **308**, 33–36.

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.

Schreiner, C.E. & Langner, G. (1997) Laminar fine structure of frequency organization in auditory midbrain. *Nature*, **388**, 383–386.

Smith, J.C., Marsh, J.T. & Brown, W.S. (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr. Clin. Neurophysiol.*, **39**, 465–472.

Sohmer, H. & Pratt, H. (1977) Identification and separation of acoustic frequency following responses (FFRs) in man. *Electroencephalogr. Clin. Neurophysiol.*, **42**(4), 493–500.

Spiegel, M.F. & Watson, C.S. (1984) Performance on frequency-discrimination tasks by musicians and nonmusicians. *J. Acoust. Soc. Am.*, **766**, 1690–1695.

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. *Eur. J. Neurosci.*, **29**, 661–668.

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. *Hear. Res.*, **261**, 22–29.

Tervaniemi, M. (2009) Musicians–same or different? *Ann. NY Acad. Sci.*, **1169**, 151–156.

Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K. & Naatanen, R. (1997) The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neurosci. Lett.*, **226**, 1–4.

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. *Exp. Brain Res.*, **161**, 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. *Eur. J. Neurosci.*, **30**, 1636–1642.

Tramo, M.J., Cariani, P.A., Delgutte, B. & Braida, L.D. (2001) Neurobiological foundations for the theory of harmony in western tonal music. *Ann. NY Acad. Sci.*, **930**, 92–116.

Wong, P.C., Skoe, E., Russo, N.M., Dees, T. & Kraus, N. (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.*, **10**, 420–422.

Zatorre, R. & McGill, J. (2005) Music, the food of neuroscience? *Nature*, **434**, 312–315.