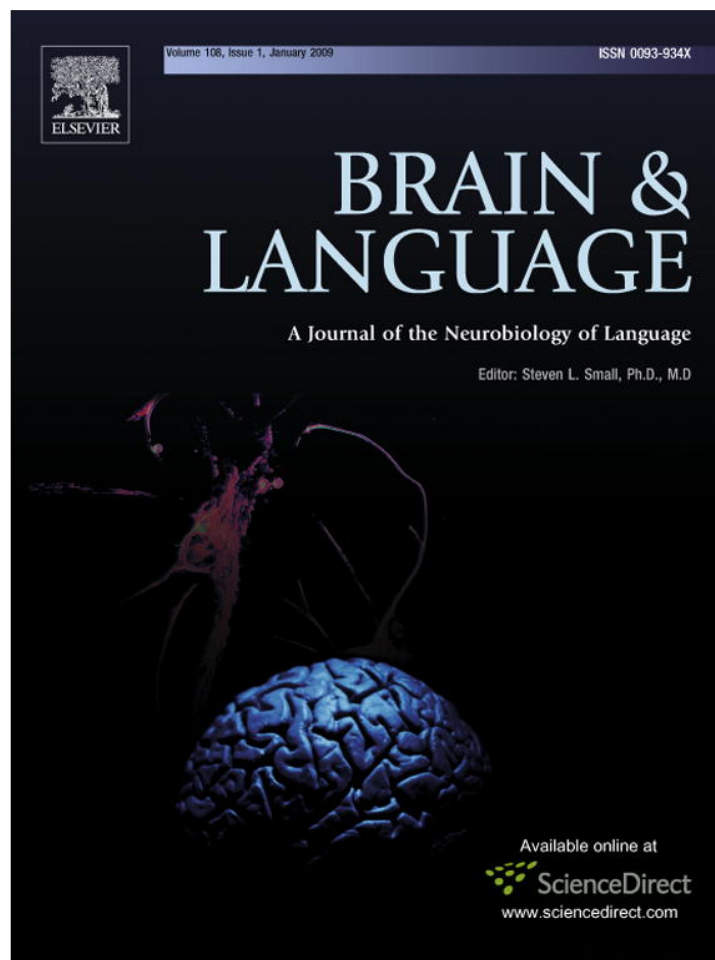


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



## Relative influence of musical and linguistic experience on early cortical processing of pitch contours

Bharath Chandrasekaran, Ananthanarayan Krishnan, Jackson T. Gandour\*

*Department of Speech Language Hearing Sciences, Purdue University, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038, USA*

Accepted 17 February 2008

Available online 17 March 2008

### Abstract

To assess domain specificity of experience-dependent pitch representation we evaluated the mismatch negativity (MMN) and discrimination judgments of English musicians, English nonmusicians, and native Chinese for pitch contours presented in a nonspeech context using a passive oddball paradigm. Stimuli consisted of homologues of Mandarin high rising (T2) and high level (T1) tones, and a linear rising ramp (T2L). One condition involved a between-category contrast (T1/T2), the other, a within-category contrast (T2L/T2). Irrespective of condition, musicians and Chinese showed larger MMN responses than nonmusicians; Chinese larger than musicians. Chinese, however, were less accurate than nonnatives in overt discrimination of T2L and T2. Taken together, these findings suggest that experience-dependent effects to pitch contours are domain-general and not driven by linguistic categories. Yet specific differences in long-term experience in pitch processing between domains (music vs. language) may lead to gradations in cortical plasticity to pitch contours. © 2008 Elsevier Inc. All rights reserved.

**Keywords:** Experience-dependent plasticity; Mismatch negativity (MMN); Music; Language; Nonspeech stimuli; Iterated rippled noise (IRN); Pitch; Lexical tone; Mandarin; Speech perception

### 1. Introduction

Both music and language exploit time-varying pitch patterns to convey information. In music, melodies are created using two types of pitch information: a contour code, involving changes in pitch direction between successive tones; and an interval code, involving the relationship between successive tones on a musical scale. In language, variation in pitch may signal prosodic contrasts at different levels of linguistic representation, i.e., syllable, word, and sentence. More than half the languages in the world are *tone languages*, so called because they exploit phonologically contrastive variations in pitch at the word or syllable level. For example, Mandarin Chinese, a contour tone language, has four tones (*ma*<sup>1</sup> ‘mother’ [T1], *ma*<sup>2</sup> ‘hemp’ [T2], *ma*<sup>3</sup> ‘horse’ [T3], *ma*<sup>4</sup> ‘scold’ [T4]), described phonetically as high level, high rising, low falling rising, and high falling,

respectively. Tone languages arguably provide an optimal window for investigating how long-term experience with time-varying pitch patterns shape perceptual and neural processing of pitch.

Pitch is a multidimensional perceptual attribute. A number of dimensions (e.g., height, direction) may serve as cues to tonal identification. The perceptual saliency of these dimensions may be influenced by the presence of specific types of pitch patterns in a language’s tonal inventory (Gandour, 1983; Gandour & Harshman, 1978) as well as by the occurrence of abstract tonal rules in the listeners’ phonological system (Hume & Johnson, 2001). Based on multidimensional scaling analysis of dissimilarity judgments, three primary dimensions appear to underlie the tone space, interpretively labeled *height*, *direction*, and *contour* (Gandour, 1983). While listeners of typologically and genetically unrelated languages seem to employ the same number of dimensions, they differ in the relative importance attached to particular dimensions. The height dimension, for example, is reported to be important across

\* Corresponding author. Fax: +1 765 494 0771.

E-mail address: [gandour@purdue.edu](mailto:gandour@purdue.edu) (J.T. Gandour).

languages, whereas the direction and contour dimensions are relatively more important to speakers of tone languages. We presume that such behavioral data reflect a relatively late attention-modulated stage of auditory processing.

The question arises as to whether pitch processing in the brain may be similarly influenced by language experience, especially at *early, preattentive* stages of cortical processing. The mismatch negativity (MMN), a frontocentrally distributed event-related potential, is an excellent tool for investigating automatic processing of suprasegmental information in speech. Using a passive oddball paradigm, the MMN has been shown to be influenced by differences in the relative saliency of pitch dimensions across languages (Chandrasekaran, Krishnan, & Gandour, 2007b). Indeed, native speakers of Mandarin show a larger MMN than English in response to exemplars of dissimilar tones (T1/T3; standard/deviant), but not in response to those that are acoustically similar (T2/T3). However, the MMN response to T1/T3 is larger than T2/T3 for the Chinese group only. Because of this interaction between stimuli, all of which are excellent phonetic exemplars of Mandarin tones, and language group, the MMN cannot simply be attributed to the activation of long-term memory traces of individual tones. A multidimensional scaling analysis of the MMN responses (Chandrasekaran, Gandour, & Krishnan, 2007) reveals two pitch dimensions (*height, contour*) across language groups. However, contour is relatively more important for Chinese than English. Thus, the MMN may serve as an index of pitch features that are differentially weighted depending on a listener's experience with lexical tones and their acoustic correlates within a given tone space. A related question is whether language-dependent neural plasticity in pitch processing is speech-specific (Chandrasekaran, Krishnan, & Gandour, 2007a). Stimuli consisted of nonspeech homologues of phonetic exemplars of Mandarin tones (T1, T2) and a nonoccurring *linear* approximation of T2. Mandarin speakers show larger MMN responses than English only to phonetic exemplars that occur in natural speech. Taken together, these findings suggest that experience-dependent neural plasticity in early preattentive cortical processing is sensitive to linguistically relevant phonetic exemplars but not specific to speech per se.

Musical experience has also been shown to modulate the preattentive processing of *musically relevant* pitch contours (Münte, Altenmüller, & Jäncke, 2002; Pantev et al., 1998; Shahin, Bosnyak, Trainor, & Roberts, 2003). As reflected by the MMN, musicians are better able to detect impure chords than nonmusicians (Koelsch, Schröger, & Tervaniemi, 1999). With respect to pitch dimensions, musicians show enhanced MMN responses to changes in the global pitch contour and pitch interval relative to nonmusicians (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004). Other factors besides musical experience per se may also influence the magnitude of the mismatch negativity. Both musicians and nonmusicians show enhanced MMN responses to

pitch changes in familiar (western) relative to nonfamiliar (nonwestern) scales (Brattico, Näätänen, & Tervaniemi, 2001). After short-term auditory training, musicians who predominantly 'learn by ear' are better able to process musical contours than classical musicians who rely on musical scores (Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001).

Given that music and language utilize complex time-varying pitch patterns, a topic of theoretical interest is whether training in the musical domain provides an advantage in the processing of *linguistically relevant* pitch contours (Patel & Iversen, 2007). Musical expertise has clearly been shown to influence neural processing of linguistic pitch contours at *attentive* stages of processing (Magne, Schön, & Besson, 2006; Marques, Moreno, Luis Castro, & Besson, 2007; Moreno & Besson, 2005; Schön, Magne, & Besson, 2004). Even at a *preattentive* stage of processing in the brainstem, musicians exhibit more accurate pitch tracking of Mandarin tones (T2, T3) than nonmusicians (Wong, Skoe, Russo, Dees, & Kraus, 2007). Since none of the musicians had previous exposure to Mandarin, their findings suggest that at least some aspects of pitch processing are domain-general.

With respect to the processing of linguistic pitch contours, previous work has focused on the effects of either musical expertise (nonnative musicians vs. nonmusicians) or linguistic expertise (native vs. nonnative speakers). To fill the gap, we must compare directly the two domains of expertise (nonnative musicians vs. native speakers). Then we can assess to what extent qualitative and quantitative differences in the processing of linguistic pitch contours vary as a function of domain (music vs. language). Contour tone languages are especially advantageous for this purpose. They employ rapid changes in pitch movement within relatively short time frames and exhibit *curvilinear* pitch trajectories. Both features have been found to be relevant in crosslanguage comparisons of early preattentive pitch processing at the level of the brainstem (Xu, Krishnan, & Gandour, 2006) and cerebral cortex (Chandrasekaran et al., 2007b). Music, on the other hand, employs comparatively slow changes in pitch movement within relatively long time frames, and moreover, is characterized by predominantly steady-state pitch.

The overall objective of this study accordingly is to determine whether processing of pitch contours, as reflected by the MMN, is specific to the domain of experience (language vs. music). The MMN is elicited using a passive oddball paradigm, eliminating task-related attention or memory confounds, and thereby representing an automatic, *preattentive* level of processing. In addition, a behavioral discrimination task allows us to examine pitch processing under conditions that involve memory demands at an *attentive* level of processing. A comparison of MMN responses from musicians and nonmusicians enables us to determine if musical training provides an advantage in the *cortical*, preattentive processing of linguistically relevant pitch contours (cf. Wong et al., 2007, brainstem).

And finally, a comparison of musicians with native speakers of Mandarin permits us to evaluate whether the duration of exposure to specific pitch contours influences the magnitude of the MMN.

To compare sensitivity of the MMN to acoustic shifts in stimuli that fall between- or within-tonal categories, three iterated rippled noise (IRN) stimuli were selected (Swaminathan, Krishnan, & Gandour, 2008), two of which represent curvilinear Mandarin tones (T1, T2) that occur in natural speech. A third was chosen to represent a linear rising ramp (T2L) that does not occur in Mandarin natural speech, or for that matter, any language of the world. T2L had the same onset, offset, and direction of pitch change as T2, but differed on the basis of overall rising trajectory (linear vs. curvilinear). Though linear ramps commonly occur in nonspeech contexts, they represent at best a crude approximation of natural speech tonal contours, and therefore are less likely to give any perceptual advantage to native speakers of a tone language. Yet it has been shown perceptually that a physical continuum of *linear*  $f_0$  trajectories ranging from level (T1) to rising (T2L) elicits language-dependent categorical effects (Xu, Gandour, & Francis, 2006).

By using IRN stimuli, the MMN of the native group cannot be attributed to a lexical–semantic bias or other psycholinguistic variables since spectrally IRN stimuli do not show any formant structure. Due to the lack of a highly modulated envelope, IRN stimuli are especially advantageous for examining pitch processing without the confound of co-varying amplitude (Yost, Patterson, & Sheft, 1996, 1998). Time-invariant IRN stimuli with constant pitch elicit robust cortical responses, both electrical (Jones, 2006) and magnetic (Krumbholz, Patterson, Seither-Preisler, Lammertmann, & Lutkenhoner, 2003; Soeta, Nakagawa, & Tonoike, 2005). Time-variant IRN stimuli with *curvilinear* pitch trajectories, representative of Mandarin tonal contours (T1/T2), elicit differential MMN responses depending on language experience (Chandrasekaran et al., 2007a). We infer that IRN curvilinear stimuli are ecologically valid inasmuch as they induce qualitatively similar experience-dependent effects as homologous *speech* stimuli (Chandrasekaran et al., 2007b).

To elicit the MMN, we constructed two passive oddball conditions: T1/T2, T2L/T2. The former involves a between-category contrast (T1/T2); the latter, a within-category contrast (T2L/T2). By including T2L/T2 we are able to evaluate whether the MMN of native speakers reflects category effects per se or more basic auditory processes. If MMN is driven by categories (Näätänen, 2001), we expect the Chinese group to show comparable profiles for electrophysiology and behavior, i.e., we expect the MMN to be larger and discrimination more accurate in T1/T2 than T2L/T2. If, on the other hand, more basic auditory processes drive the MMN, we predict larger responses in the Chinese group across conditions, relative to musicians, regardless of any interactions between group (native vs. nonnative) and stimuli (between- vs. within-category) on a discrimination task. This experimental outcome would

suggest that native MMN responses to linguistically relevant pitch contours are enhanced due to the context of their long-term learning experience. In the case of musicians, they are expected to show larger MMN responses and better discrimination than nonmusicians across conditions (T1/T2, T2L/T2). This experimental outcome would support the view that experience-dependent plasticity to linguistically relevant pitch contours is not specific to the language domain.

## 2. Materials and methods

### 2.1. Participants

Eleven adult native speakers of Mandarin (5 men, 6 women), 11 adult native speakers of American English with musical training (5 men, 6 women), and 11 adult native speakers of American English without musical training (5 men, 6 women) took part in the ERP and behavior experiments. The three groups were closely matched in age (Mandarin:  $M = 26.4$  years,  $SD = 4.2$ ; musicians:  $M = 24.2$ ,  $SD = 3.2$ ; nonmusicians:  $M = 25.2$ ,  $SD = 3.5$ ), education (Mandarin:  $M = 18.2$ ,  $SD = 2.2$ ; musicians:  $M = 16.2$ ,  $SD = 2.2$ ; nonmusicians:  $M = 17.4$ ,  $SD = 1.6$ ) and were strongly right handed (90%) as measured by the Edinburgh Handedness inventory (Oldfield, 1971). At the time of their participation, all participants were graduate students at Purdue University, West Lafayette, Indiana. All participants exhibited normal hearing sensitivity (20 dB HL) at frequencies of 0.5, 1, 2, and 4 kHz. In addition, participants reported no previous history of neurological or psychiatric illnesses. All participants completed a language history questionnaire (Li, Sepanski, & Zhao, 2006) and a music history questionnaire (Wong & Perrachione, 2007). Native speakers of Mandarin were all originally from mainland China, and did not have any formal English instruction before the age of 11 years ( $M = 12.5$ ,  $SD = 1.5$ ) in secondary school. The two American English groups had no prior experience with learning any tonal language, and were not exposed to any second language before the age of 11 years. Musically trained participants were amateur instrumentalists who had at least 8 years of continuous

Table 1  
Musical background of amateur musicians

Participant	Instrument	Years of training	Age of onset
M1	Piano/violin	17	2
M2	Violin	08	5
M3	Piano	13	4
M4	Viola	12	7
M5	Piano/flute	11	9
M6	Piano	12	5
M7	Piano	12	6
M8	Violin	13	6
M9	Piano	11	7
M10	Piano	14	6
M11	Piano/guitar/percussion	13	7

training on an instrument ( $M = 12.36$ ,  $SD = 2.20$ ), starting at or before the age of 10 years ( $M = 6$ ,  $SD = 1.95$ ) (Table 1). In contrast, none of the native Mandarin or musically untrained American English participants had more than 3 years of formal training in music or any combination of instruments, and none had any musical training within the past 5 years. They were paid for their participation. They gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

## 2.2. Stimuli

Iterated rippled noise (IRN) was used to create three time-varying nonspeech  $f_0$  contours using procedures similar to those described in Swaminathan et al. (2008). A high iteration step (32) was used for all contrasts with gain set to 1. At high iteration steps, clear bands of energy are present at the  $f_0$  and its harmonics, but unlike speech, IRN stimuli show no formant structure or temporal envelope. Of the three time-varying  $f_0$  contours (Fig. 1), two (T1, T2) were modeled after natural citation-form Mandarin  $f_0$  contours using a fourth-order polynomial equation (Xu, 1997). T1 and T2 reflect native Mandarin high level and rising tones, respectively, differing from each other on the basis of onset, offset, height, direction, and shape of  $f_0$  contour. The third stimulus (T2L) was a linear approximation of T2, having the same onset, offset, and direction as T2, but differing from T2 on the basis of its trajectory, i.e., linear vs. curvilinear. The duration of all three stimuli was fixed at 250 ms. Amplitude was fixed at 70 dB. Ten milliseconds of cosine

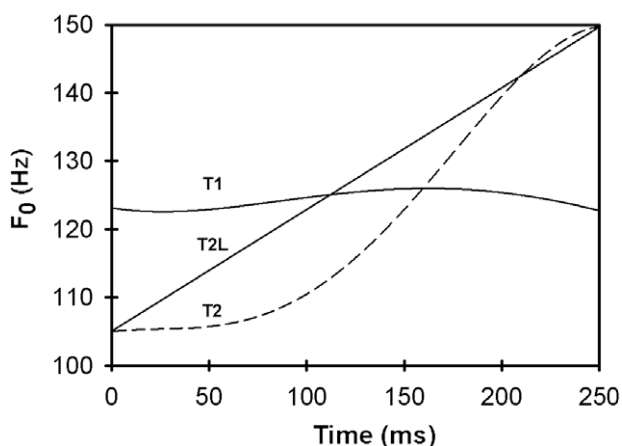


Fig. 1. Voice fundamental frequency contours of the three IRN stimuli. T1 and T2 are modeled after average  $f_0$  contours of time-normalized Mandarin lexical tones (Swaminathan et al., 2008). T2L, a linear rising ramp, represents a pitch contour that, albeit a crude approximation of T2, does not actually occur in natural speech. T2 (dashed line) was the deviant in both ERP conditions. T1 was the standard in one condition (T1/T2); T2L in the other (T2L/T2). T2 and T2L have identical onsets, offsets, and direction of  $f_0$  movement (rising). They differ primarily on the basis of contour; T2 is curvilinear, whereas T2L is linear. T1, Mandarin high level tone; T2, Mandarin high rising tone; T2L, linear rising ramp that does not occur in Mandarin tonal inventory.

rising and falling ramps was added to reduce spectral splatter.

## 2.3. ERP experiment

### 2.3.1. Data acquisition

Participants sat in an acoustically and electrically shielded chamber facing a LCD monitor that was connected to a DVD player. Participants were instructed to ignore the sounds presented binaurally via insert earphones and watch a self-selected movie with subtitles. In order to ensure that attention was focused on the movie, and not the aural presentation, participants were informed that they would have to provide a synopsis of the movie at the end of the experimental session. The interstimulus onset-to-onset interval was fixed at 667 ms. For all oddball sequences, the frequent stimulus (standard) was presented at a probability of 0.85 and the infrequent stimulus (deviant) occurred at a probability of 0.15. Within oddball sequences, the order of presentation of stimuli was pseudo-random, i.e., at least one standard stimulus preceded the deviant.

The experiment consisted of four oddball sequences. In one condition (T1/T2), T1 (curvilinear: level) was presented as the standard ( $p = 0.85$ ) and T2 (curvilinear: rising) as the deviant ( $p = 0.15$ ). In a second condition (T2L/T2), T2L (linear: rising) was presented as the standard, with T2 (curvilinear: rising) as the deviant. Thus, in these two oddball sequences, T2 was the common deviant, occurring either in the context of T1 or T2L. In the two remaining sequences, the oddball sequences were reversed with T2 as the standard ( $p = 0.85$ ) with T1 or T2L occurring as the deviants ( $p = 0.15$ ). Hundred artifact-free deviants were collected from each sequence. The experiment ran for approximately 2 h including subject preparation. All stimuli were controlled by a signal generation and data acquisition system with a 4-channel optical differential amplifier (Intelligent Hearing Systems, Smart EP). Stimuli were presented binaurally at 75 dB SPL through magnetically shielded insert earphones (Biologic TIP-300).

For each participant, AgCl electrodes were mounted on the frontal midline (Fz), central midline (Cz) locations according to the 10–20 location system. These two electrode locations were chosen because the typical MMN response is known to be the most robust at the frontal electrode sites (Näätänen et al., 1997) and shows a distinct reduction in amplitude at more central sites. The tip of the nose served as the reference electrode, and the forehead served as the ground. The right and left mastoids were linked and used as a third reference site. Since the MMN is known to invert at the mastoid electrodes, we were able to confirm if the negativity is a true MMN by examining this reference site. The impedance across all electrodes was kept below 5 k $\Omega$ . Electrodes monitoring vertical eye movements were used to remove eye-blink-related artifacts. Epochs with voltage changes exceeding 60  $\mu$ V were automatically removed

online. The signals were band-pass filtered at 1–30 Hz and recorded at a 1000 Hz sampling rate.

### 2.3.2. Data analysis

The baseline for the grand averaged waveforms was defined as the average of the amplitude values between –100 and 0 ms (onset of stimuli). To obtain the MMN, the deviant waveforms from the T1/T2 and T2L/T2 sequences were subtracted from standard waveforms presented, respectively, in the T2/T1 and T2/T2L sequences. Subtracting the deviant from the same stimuli presented as the standard effectively controls for any acoustical differences between stimuli. The MMN peak latency was calculated as the most negative voltage in the MMN window between 125 and 300 ms. The MMN mean amplitude was calculated as the mean voltage from a 100 ms window centered on the MMN peak latency.

The MMN mean amplitude and peak latencies were analyzed using a three-way mixed model ANOVA for the effects of group (Chinese, musicians, nonmusicians), condition (T1/T2, T2L/T2), and location (Fz, Cz). In the ANOVA model, subjects (random) were nested within group, a between-subject factor; condition and location (fixed) were within-subject factors.

### 2.4. Behavioral experiment

#### 2.4.1. Data acquisition

Subjects were asked to perform speeded response discrimination judgments of the IRN nonspeech stimuli (T1, T2, T2L) immediately following the ERP experiment. They were first presented with a practice set of stimuli in order to gain familiarity with the task. Each trial consisted of a pair of stimuli including a 300 ms interstimulus interval. The ‘same’ and ‘different’ trials had equal probability of occurrence. All trials were randomized within each block. The two stimuli within ‘different’ trials were also presented in random order. Subjects were asked to press the left (‘same’) or right (‘different’) mouse button to indicate their discrimination judgment during the 1.5 s response interval following each pair. Stimuli were presented binaurally by means of computer playback (E-Prime) through a pair of Sony MDR-7506 headphones at a comfortable listening level (72 dB SPL).

#### 2.4.2. Data analysis

Response accuracy (%) was calculated for each subject and tone pair (T1/T2, T2L/T2). The ‘different’ pairs were pooled across order of presentation of tones within pairs, yielding 20 trials per condition (T1/T2, T2L/T2). Arcsine-transformed proportions of correct responses (Winer, Brown, & Michels, 1991) were subjected to a mixed model ANOVA for the effects of group (between-subjects: Chinese, musicians, nonmusicians) and condition (within-subjects: T1/T2, T2L/T2) to determine whether discrimination accuracy varied as a function of language experience.

## 3. Results

### 3.1. MMN morphology

The grand average waveforms for the three groups (Chinese, musicians, nonmusicians), two conditions (T1/T2, T2L/T2), and three locations (Fz, Cz, linked mastoids) are shown in Fig. 2. Both conditions (T1/T2, T2L/T2) elicited robust MMN responses for all three groups within the 125–300 ms time window. The MMN reduced in amplitude at Cz, relative to Fz, and reversed in polarity at the mastoid location, indicative of a ‘true’ mismatch response.

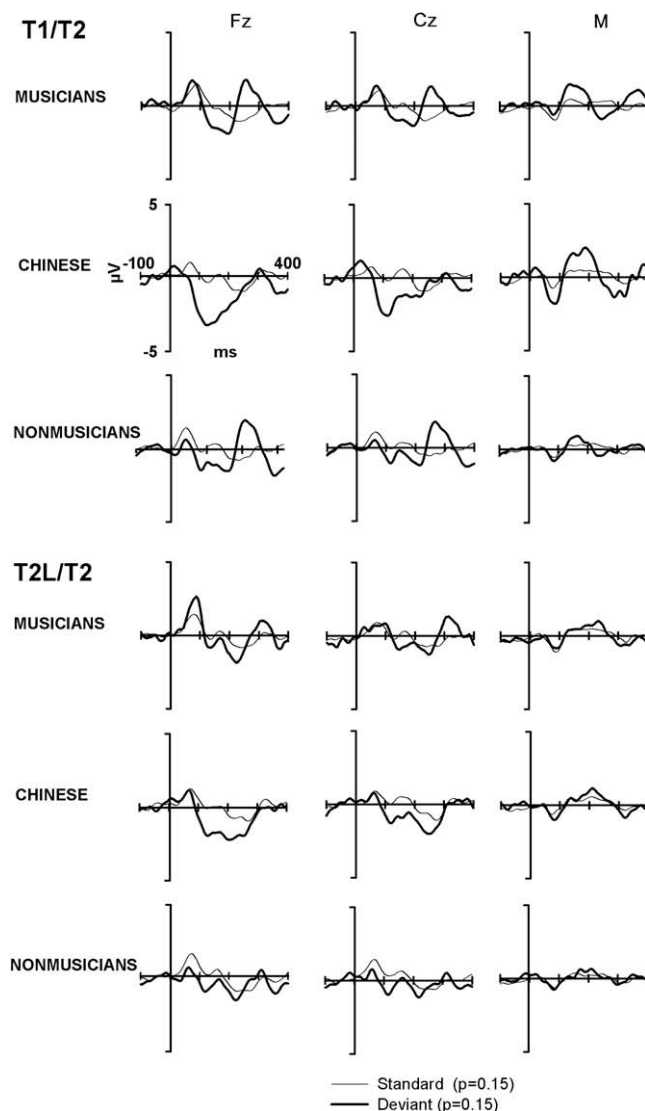


Fig. 2. Grand average standard ( $p = 0.85$ ) and deviant ( $p = 0.15$ ) waveforms per group (Chinese, musicians, nonmusicians) and condition (T1/T2, T2L/T2) at three electrode locations (Fz, Cz, mastoid). Irrespective of group or condition, the MMN was larger at Fz than at Cz, and showed the typical polarity reversal at the mastoids. With respect to group, MMN-related negativity was larger for musicians relative to nonmusicians across conditions. Chinese subjects, on the other hand, showed a larger MMN response than either musicians or nonmusicians. For all groups, the peak latency of the MMN was later in the T2L/T2 than in the T1/T2 condition.

### 3.2. MMN mean amplitude

The grand average waveforms for the three groups (Chinese, musicians, nonmusicians), two conditions (T1/T2, T2L/T2), and three locations (Fz, Cz, linked mastoids) are shown in Fig. 2. Results from the omnibus three-way ANOVA (group  $\times$  condition  $\times$  location) revealed significant main effects of group [ $F_{2,30} = 36.10$ ;  $p < 0.0001$ ,  $\eta^2_{\text{partial}} = 0.71$ ], condition [ $F_{1,30} = 9.36$ ;  $p = 0.004$ ,  $\eta^2_{\text{partial}} = 0.24$ ] and location [ $F_{1,60} = 39.22$ ;  $p < 0.0001$ ,  $\eta^2_{\text{partial}} = 0.40$ ]. There were no significant two- or three-way interaction effects between group, condition, and location.

The MMN mean amplitude for each group (Chinese, musicians, nonmusicians) and condition (T1/T2, T2L/T2) at the electrode location Fz are displayed in the left panel of Fig. 3. Pooling across conditions, post hoc Tukey–Kramer adjusted comparisons revealed that the Chinese group had a larger MMN mean amplitude than either musicians [ $t_{30} = 4.14$ ,  $p = 0.0007$ ] or nonmusicians [ $t_{30} = 8.50$ ,  $p < 0.0001$ ]. Musicians, in turn, had a larger MMN mean amplitude relative to nonmusicians [ $t_{30} = 4.36$ ,  $p = 0.0004$ ]. Pooling across groups, the MMN mean amplitude of the T1/T2 condition was significantly greater than T2L/T2 [ $t_{30} = 3.06$ ,  $p = 0.004$ ].

### 3.3. MMN peak latency

The mean peak latency for each group and condition at the electrode location Fz is plotted in the right panel of Fig. 3. Results from the omnibus ANOVA yielded a significant main effect of condition [ $F_{1,30} = 25.47$ ;  $p < 0.0001$ ,  $\eta^2_{\text{partial}} = 0.46$ ]. The main effect of group or location failed to reach significance. None of the two- or three-way inter-

actions reached significance. Pooling across groups, MMN peak latency of T2L/T2 was significantly later than T1/T2 [ $t_{30} = 5.05$ ,  $p < 0.0001$ ].

### 3.4. Behavioral performance

A two-way (group  $\times$  condition) repeated measures ANOVA conducted on the arcsine-transformed proportion of correct responses yielded significant main effects of group [ $F_{2,30} = 10.62$ ;  $p = 0.0003$ ,  $\eta^2_{\text{partial}} = 0.41$ ] and condition [ $F_{1,30} = 18.32$ ;  $p = 0.0002$ ,  $\eta^2_{\text{partial}} = 0.40$ ], and a significant interaction effect between group and condition [ $F_{2,30} = 10.36$ ;  $p = 0.0004$ ,  $\eta^2_{\text{partial}} = 0.41$ ]. Per group, only Chinese subjects performed at a higher level of accuracy in the T1/T2 condition as compared to T2L/T2 [ $F_{1,30} = 38.26$ ;  $p < 0.0001$ ]; cf. musicians [ $F_{1,30} = 0.29$ ;  $p = 0.60$ ], nonmusicians [ $F_{1,30} = 0.48$ ;  $p = 0.49$ ]. Per condition, only T2L/T2 yielded a group difference in response accuracy [ $F_{2,30} = 20.98$ ;  $p < 0.0001$ ]. In the T2L/T2 condition, the Chinese group was less accurate relative to either musicians [ $t_{30} = -5.68$ ,  $p < 0.0001$ ] or nonmusicians [ $t_{30} = -5.53$ ,  $p < 0.0001$ ] (Fig. 4).

## 4. Discussion

### 4.1. Modulation of the mismatch negativity to linguistic pitch is not circumscribed to the language domain

In this study, we report that with respect to early, preattentive cortical responses to IRN homologues of linguistically relevant pitch patterns, native Mandarin speakers show larger MMN responses than nonnative English-speaking amateur musicians or nonmusicians in both odd-

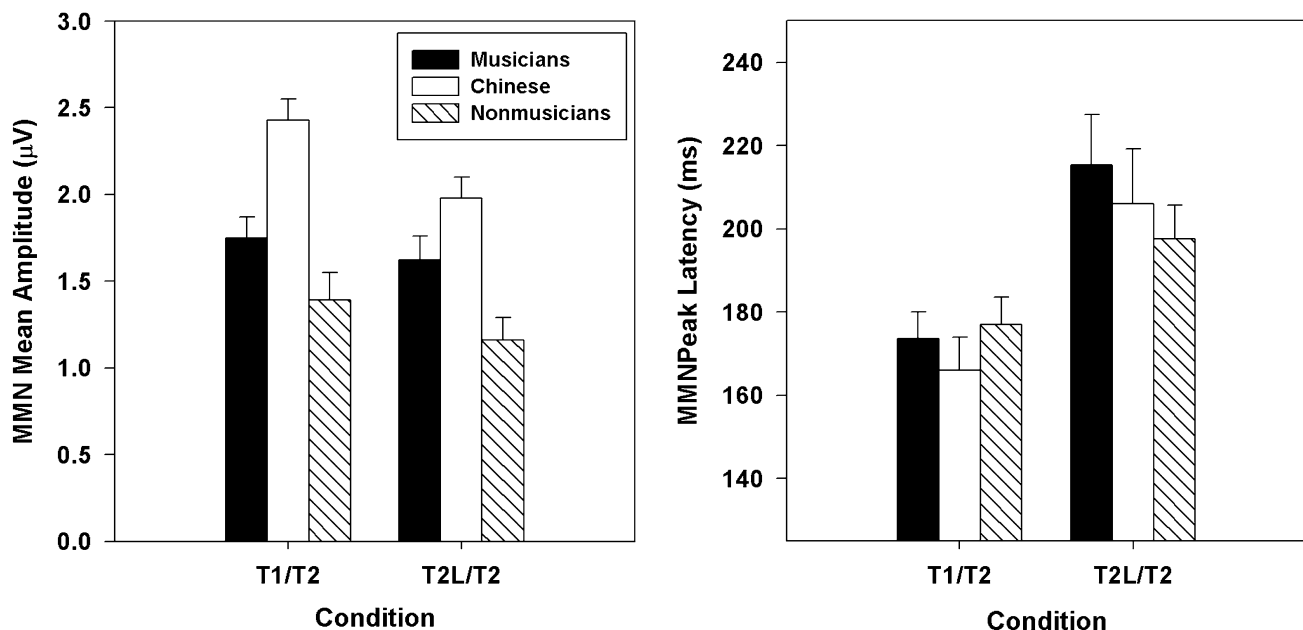


Fig. 3. Mean MMN amplitude (left panel) and peak latency (right panel) values for the three groups (Chinese, musicians, nonmusicians) per condition (T1/T2, T2L/T2) as measured from the Fz electrode location. Regardless of condition, the mean MMN amplitude, in order of groups from largest to smallest, was Chinese, musicians, and nonmusicians. Irrespective of group, the MMN peaked later for T2L/T2 relative to T1/T2.

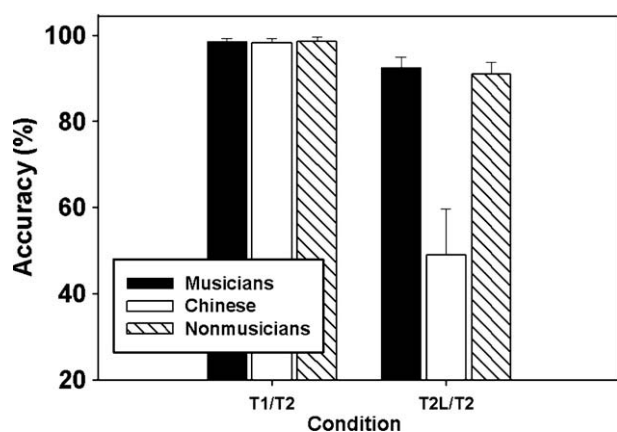


Fig. 4. Discrimination accuracy, as a function of paired  $f_0$  contours (T1/T2, T2L/T2) and group (Chinese, musicians, nonmusicians). All groups performed at ceiling level for T1/T2. For T2L/T2, Chinese subjects were less accurate than either native English musicians or nonmusicians. Only the Chinese group was significantly less accurate for T2L/T2 relative to T1/T2.

ball sequences (T1/T2; T2L/T2). This finding demonstrates that experience-dependent plasticity of pitch processing is not speech-specific, but is sensitive to the context of the long-term experience (native vs. nonnative). We further report that musicians, in turn, exhibit larger MMN responses than nonmusicians across sequences. This finding demonstrates that experience-dependent plasticity in early, preattentive cortical processing is not circumscribed to the domain (language vs. music) in which the pitch patterns are of behavioral relevance.

It has been proposed that crosslanguage differences in the mismatch negativity reflect the influence of phonetic categories (Näätänen et al., 1997; Sharma & Dorman, 2000). The observation that MMN responses are larger for native relative to nonnative speakers is attributed to the presence of long-term stored representations of such categories. In this study, we included a nonnative group who had long-term experience with nonlinguistic pitch patterns (music) but no prior experience with a tonal language. It is safe to assume that English musicians do not have long-term representations of Mandarin tonal categories. The fact that MMN responses are larger for musicians than for nonmusicians suggests that an explanation based exclusively on the presence or absence of categories is untenable.

Neither can phonetic categories be invoked for explaining experience-dependent plasticity for the native Chinese group. Behaviorally, Chinese subjects were *less* accurate than either musicians or nonmusicians in discrimination of T2L vs. T2. Furthermore, only the Chinese group was less accurate in discrimination of T2L vs. T2 as compared to T1 vs. T2. These data suggest that Chinese speakers judged T2L and T2 to be from the same category, and are consistent with a previous study of categorical perception of Mandarin tones using *linear rising ramps* to represent T2 (Xu et al., 2006). On attention-demanding tasks, behavioral or otherwise, we conclude that native perception of pitch patterns is modulated by long-term stored rep-

resentations of tonal categories. If the MMN to linguistic pitch patterns was similarly driven by tonal categories, we would expect less robust responses for the T2L/T2 condition relative to the T1/T2 condition for the Chinese group. Contrary to fact, we find that among the three groups, Chinese subjects show the largest MMN responses for the T2L/T2 condition.

So how do we account for experience-dependent plasticity in early cortical processing of linguistic pitch patterns for *both* native speakers and musicians? As measured by frequency-following responses (FFRs), a neural ensemble response reflecting brainstem phase-locking, it has been demonstrated that both groups exhibit stronger representations of linguistic pitch relative to nonnative nonmusicians at the level of the brainstem (Krishnan, Xu, Gandour, & Cariani, 2005; Wong et al., 2007). These brainstem data suggest that both native speakers of Mandarin as well as nonnative musicians receive more robust representations of pitch relevant information to the auditory change-detection process underlying the MMN. With stronger pitch representation of the 'standard' and 'deviant' traces, this process is likely to be more efficient for native speakers and musicians as compared to musically untrained speakers of English.

#### 4.2. Mismatch negativity is sensitive to the context of learning experience

Yet we observe that the *degree* of MMN magnitude differs between musicians and native Chinese speakers. Native Chinese show more robust MMN responses than musicians irrespective of condition. This finding suggests that the representation of linguistically relevant pitch patterns is stronger for the Chinese group than musicians at early stages of cortical processing, presumably reflecting differences in the context of learning experience (music vs. language). Indeed, cortical plasticity has been shown to be highly sensitive to the context of learning (Gilbert, Sigman, & Crist, 2001) as well as the behavioral relevance of the stimuli (Polley, Steinberg, & Merzenich, 2006; Rutkowski & Weinberger, 2005; Weinberger, 2004).

As a result of their long-term experience in listening to time-varying pitch contours that are linguistically relevant at the syllable level, the Chinese group's MMN responses may index processing strategies that have been streamlined for dynamic pitch changes over relatively short time intervals. In tone languages, fairly rapid  $f_0$  movements are required for high intelligibility of contour tones, whether rising or falling (Abramson, 1978). Chinese subjects are more accurate than English subjects in identifying changes in pitch direction of nonspeech FM (frequency-modulated) sweeps, even at time scales much shorter than those normally associated with lexical tones (Luo, Boemio, Gordon, & Poeppel, 2007). Chinese subjects exhibit larger MMN responses than English subjects when processing linguistically relevant pitch contours (Chandrasekaran et al., 2007b), and furthermore show greater sensitivity to fine-



grained changes in the shape of pitch contours (Chandrasekaran et al., 2007a). Speech production data also show that  $f_0$  patterns of Mandarin have a greater amount of dynamic movement as a function of time and number of syllables than those of English (Eady, 1982). Taken as a whole, the extant literature reveals that native speakers of a tone language have a built-in advantage in processing rapidly changing pitch contours.

Even though these pitch trajectories are of no behavioral relevance to nonnative speakers of Mandarin, musicians show more robust early cortical processing of linguistic pitch patterns relative to nonmusicians. In visual perceptual learning, it has similarly been demonstrated that the brain can adapt to features that are not behaviorally relevant, as long as these features are presented frequently (Watanabe, Nanez, & Sasaki, 2001). It is possible that musicians, who are experts in pitch processing related to music, adapt to the features of linguistic pitch patterns over the time course of stimulus presentation. At the level of the brainstem (Wong et al., 2007), musicians show superior pitch tracking accuracy of linguistic pitch patterns as compared to nonmusicians. Furthermore, these authors find a high positive correlation between pitch tracking accuracy and number of years of musical training suggesting that enhanced pitch tracking is related to their experience in pitch processing in the *music domain*.

Our data converge with previous electrophysiological studies that have shown enhanced neural processing of pitch for musicians relative to nonmusicians (Fujioka et al., 2004; Koelsch et al., 1999; Kuriki, Kanda, & Hirata, 2006; Shahin, Roberts, & Trainor, 2004). Since musicians show larger MMN responses to nonspeech homologues of Mandarin tones (T1, T2), we infer that musical expertise facilitates the processing of pitch variations not only in music but also language (Marques et al., 2007; Moreno & Besson, 2005; Patel & Iversen, 2007). In cortical as well as brainstem processing of linguistic pitch patterns, musicians may have an advantage over nonmusicians in their learning of a tonal language (Patel & Iversen, 2007; Wong & Perrachione, 2007).

The exact neurobiological mechanism underlying experience-dependent enhanced tuning of the mismatch negativity to behaviorally relevant sound dimensions has not been clearly elucidated in humans. In animal models, however, a corticofugal mechanism has been hypothesized to mediate cortical neuroplasticity (Suga, Gao, Zhang, Ma, & Olsen, 2000; Suga, Ma, Gao, Sakai, & Chowdhury, 2003; Suga, Xiao, Ma, & Ji, 2002). According to Suga and colleagues, the cortex shapes the brainstem processing to repetitive sounds via the corticofugal feedback mechanism. This *short-term plasticity* improves the response properties of the auditory cortex to the incoming stimulus stream regardless of its behavioral relevance. Once the stimuli achieve behavioral relevance, there is an increase in corticofugal feedback, resulting in more enhanced subcortical tuning, and consequently *long-term cortical plasticity*. In this study, it is possible that musicians' long-term

training with dynamic, musical pitch patterns may transfer to behaviorally irrelevant pitch patterns in the language domain, a result of short-term plasticity induced by corticofugal mechanisms (Wong et al., 2007). Native Mandarin speakers, on the other hand, are likely to benefit from long-term cortical plasticity, i.e., enhanced corticofugal feedback *and/or* reorganization of cortical and brainstem circuitry resulting from the behavioral relevance of the pitch patterns. Such differences in neural mechanisms underlying cortical plasticity may explain the graded differences we observe in MMN responses between groups (Chinese > musicians > nonmusicians).

## 5. Conclusion

Both music and language experience modulate automatic early cortical processing of dynamic nonspeech pitch trajectories. Although modulation of the MMN is found to be domain-general, amateur musicians show less robust MMN responses when compared to native Mandarin speakers. Thus, domain specificity can influence the *degree* of modulation of MMN responses. Yet the Chinese group was less accurate than nonnative groups in discriminating two pitch trajectories that fall within the range of a tonal category (T2). Thus, experience-dependent plasticity at early *preattentive* stages of processing is not specific to speech or domain of experience, but is sensitive to the context of the learning experience (music vs. language). In contrast, at later attentive stages of processing, experience-driven effects are highly sensitive to categorical representations.

## Acknowledgments

This article is based on part of a doctoral dissertation to be submitted by the first author at Purdue University in May 2008. B.C. is currently a predoctoral student in the Purdue University Life Sciences Integrative Neuroscience Program. Research supported in part by the National Institutes of Health R01 DC008549-01 (A.K.) and research incentive grants from the College of Liberal Arts (A.K., J.G.). Thanks to Jayaganesh Swaminathan for his assistance in generating the IRN stimuli; Bruce Craig and Eunjung Lim for their help with statistical analysis; Edward Bartlett for his useful comments on earlier versions of the manuscript.

## References

- Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech*, 21, 319–325.
- Brattico, E., Näätänen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception*, 19(2), 199–222.
- 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, 195–210.

- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007a). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Neuroreport*, 18(18), 1963–1967.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007b). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, 1128(1), 148–156.
- Eady, S. J. (1982). Differences in the F0 patterns of speech: Tone language versus stress language. *Language and Speech*, 25(1), 29–42.
- 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.
- Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149–175.
- Gandour, J., & Harshman, R. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech*, 21, 1–33.
- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5), 681–697.
- Hume, E., & Johnson, K. (2001). A model of the interplay of speech perception and phonology. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 3–25). New York: Academic Press.
- Jones, S. J. (2006). Cortical processing of quasi-periodic versus random noise sounds. *Hearing Research*, 221(1–2), 65–72.
- Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309–1313.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research Cognitive Brain Research*, 25(1), 161–168.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13(7), 765–772.
- Kuriki, S., Kanda, S., & Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *Journal of Neuroscience*, 26(15), 4046–4053.
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavioral Research Methods*, 38(2), 202–210.
- Luo, H., Boemio, A., Gordon, M., & Poeppel, D. (2007). The perception of FM sweeps by Chinese and English listeners. *Hearing Research*, 224(1–2), 75–83.
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211.
- Marques, C., Moreno, S., Luis Castro, S., & 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.
- 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.
- Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3(6), 473–478.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38(1), 1–21.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432–434.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- 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.
- Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, 11(9), 369–372.
- Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience*, 26(18), 4970–4982.
- Rutkowski, R. G., & Weinberger, N. M. (2005). Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proceedings of the National Academy of Sciences*, 102(38), 13664–13669.
- Schön, 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.
- Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, 23(13), 5545–5552.
- Shahin, A., Roberts, L. E., & Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport*, 15(12), 1917–1921.
- Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America*, 107(5), 2697–2703.
- Soeta, Y., Nakagawa, S., & Tonoike, M. (2005). Auditory evoked magnetic fields in relation to iterated rippled noise. *Hearing Research*, 205(1–2), 256–261.
- Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: Recent progress. *Proceedings of the National Academy of Sciences*, 97(22), 11807–11814.
- Suga, N., Ma, X., Gao, E., Sakai, M., & Chowdhury, S. A. (2003). Descending system and plasticity for auditory signal processing: Neuroethological data for speech scientists. *Speech Communication*, 41, 189–200.
- Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, 36(1), 9–18.
- Swaminathan, J., Krishnan, A., & Gandour, J. T. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Transactions on Biomedical Engineering*, 55(1), 281–287.
- Tervaniemi, M., Rytönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, 8(5), 295–300.
- Watanabe, T., Nanez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, 413(6858), 844–848.
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience*, 5(4), 279–290.
- Winer, B. J., Brown, R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.
- Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565–585.
- 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. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61–83.
- Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, 120(2), 1063–1074.
- Xu, Y., Krishnan, A., & Gandour, J. T. (2006). Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport*, 17(15), 1601–1605.
- Yost, W. A., Patterson, R., & Sheft, S. (1996). A time domain description for the pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America*, 99(2), 1066–1078.
- Yost, W. A., Patterson, R., & Sheft, S. (1998). The role of the envelope in processing iterated rippled noise. *Journal of the Acoustical Society of America*, 104(4), 2349–2361.