

# Influence of Musical Expertise on Segmental and Tonal Processing in Mandarin Chinese

Céline Marie<sup>1,2</sup>, Franco Delogu<sup>3,4,5</sup>, Giulia Lampis<sup>3,4,5</sup>,  
Marta Olivetti Belardinelli<sup>3,4</sup>, and Mireille Besson<sup>1</sup>

## Abstract

■ A same–different task was used to test the hypothesis that musical expertise improves the discrimination of tonal and segmental (consonant, vowel) variations in a tone language, Mandarin Chinese. Two four-word sequences (prime and target) were presented to French musicians and nonmusicians unfamiliar with Mandarin, and event-related brain potentials were recorded. Musicians detected both tonal and segmental variations more accurately than nonmusicians. Moreover, tonal variations were associated with

higher error rate than segmental variations and elicited an increased N2/N3 component that developed 100 msec earlier in musicians than in nonmusicians. Finally, musicians also showed enhanced P3b components to both tonal and segmental variations. These results clearly show that musical expertise influenced the perceptual processing as well as the categorization of linguistic contrasts in a foreign language. They show positive music-to-language transfer effects and open new perspectives for the learning of tone languages. ■

## INTRODUCTION

Asian tone languages such as Mandarin, Cantonese, or Thai have recently been the focus of increased research interests for at least two reasons. In contrast to nontone languages such as English or French, pitch variations are linguistically relevant and determine the meaning of words (Xu, 2001; Xu & Wang, 2001). In Mandarin, for example, there are four contrastive tones: Tone 1 is high-level; Tone 2 is high-rising; Tone 3 is low-dipping; and Tone 4 is high-falling (i.e., /di/ with high-level Tone 1 means “low,” but /di/ with high-falling Tone 4 means “ground”; Chao, 1947). Tone languages are therefore ideally suited to examine the influence of experience-dependent linguistic knowledge on pitch processing. Typically, results of experiments using PET or fMRI have demonstrated left-lateralized pitch processing in tone language speakers (i.e., when pitch variations are linguistically meaningful) and right hemisphere lateralization in nontone language speakers (Klein, Zatorre, Milner, & Zhao, 2001; Gandour et al., 2000; Gandour, Wong, & Hutchins, 1998; and see Zatorre & Gandour, 2008 for a review).

The second reason for interest in tone languages is that they allow testing for the influence of musical expertise on linguistic pitch processing. Previous results have shown that musicians are more sensitive to subtle pitch variations than nonmusicians (e.g., Micheyl, Delhommeau,

Perrot, & Oxenham, 2006; Schön, Magne, & Besson, 2004). For instance, Micheyl et al. (2006) reported that discrimination thresholds for pure and complex harmonic tones were six times lower for musicians than for nonmusicians. The question we address here is whether this ability extends to pitch processing in speech. If the perception of pitch in music and speech relies on common mechanisms, musicians should be more sensitive to linguistic pitch variations in tone languages than nonmusicians. Such an outcome, showing that some aspects of speech processing (here pitch) are dependent upon the level of expertise in another domain (here music), would argue against the idea that the language system is “informationally encapsulated” (Fodor, 1983) and that the computations necessary to process language unfold independently of other types of knowledge (Pinker, 1997, 2005 and see Besson & Schön, in press; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Fodor, 2000, for further discussion). From a theoretical perspective, such results (showing that the sensitivity to musical pitch extends to linguistic pitch) would argue against the existence of a domain-specific linguistic pitch (sub)module that is immune to knowledge in another domain (i.e., musical expertise). From an applied perspective, they would demonstrate that musical expertise may facilitate the learning of tone languages that are widely spoken in Asia and in Africa.

Results of empirical studies using behavioral methods clearly provide some evidence in favor of the influence of musical expertise on linguistic pitch processing. Gottfried and Riester (2000) showed that English music majors with no experience of tone languages identified the four

<sup>1</sup>CNRS—Marseille Universities, France, <sup>2</sup>McMaster University, Hamilton, ON, Canada, <sup>3</sup>Sapienza University of Rome, Italy, <sup>4</sup>Inter-university Centre for Research on Cognitive Processing in Natural and Artificial Systems, Italy, <sup>5</sup>RIKEN Brain Science Institute, Wako, Japan

Mandarin tones better than nonmusicians, and that perception and production of Mandarin tones was enhanced in musicians compared to nonmusicians (Gottfried, Staby, & Ziemer, 2004; see also Alexander, Wong, & Bradlow, 2005 for similar results). Very recently, Lee and Hung (2008) used intact and modified Mandarin syllables with the four Mandarin tones produced by multiple speakers in an attempt to specify the attributes of pitch (pitch height, pitch contour, pitch variability) that English musicians (with 15 years of musical training on average) perceived better than nonmusicians. They found that musicians processed pitch contour better than nonmusicians (which was not due to absolute pitch abilities as none of the musicians had absolute pitch). Taken together, these behavioral data concur in showing lexical tone processing advantages in musicians compared to nonmusicians.

Of particular interest for the present study, Delogu, Lampis, and Olivetti Belardinelli (2006, 2010) asked Italian speakers, with no knowledge of tone languages, to perform a same-different task on sequences of monosyllabic Mandarin words. In both adults and children, results showed that melodic abilities and musical expertise enhanced the discrimination of lexical tones. However, no effect was found on the discrimination of segmental variations such as consonant or vowel changes within a word. This result stands in contrast with other results showing that musical expertise influences phoneme processing in one's own language. For instance, Slevc and Miyake (2006) found that musical ability predicts phonological ability in adult English speakers. Consistently, Anvari, Trainor, Woodside, and Levy (2002) found positive relationships between musical skills, phonological processing, and early reading ability in English preschool children. Moreover, results of a recent longitudinal study have demonstrated that 6 months of musical training in 8-year-old nonmusician children improved reading of phonologically complex words (Moreno et al., 2009). Finally, deficits in musical pitch recognition have been shown to be associated with deficits in the phonological processing of speech sounds (Jones, Lucker, Zalewski, Brewer, & Drayna, 2009).

Based on such contradictory evidence, the present study aimed at further examining the influence of musical expertise on both lexical tone and segmental processing. French musicians and nonmusicians listened to two sequences of four Mandarin monosyllabic words to decide if they were the same or different. When different, one word in the sequence was different at a tonal or at a segmental level. Based on the results reviewed above, we predicted better tone discrimination (lower percentage of errors and faster RTs) in musicians than in nonmusicians. At issue is whether musicians would also show better segmental discrimination.

Moreover, to access the temporal dynamics of brain activity related to segmental and linguistic pitch processing, we used the ERP method. In line with behavioral data, results of previous experiments using the ERPs or

the magnetoencephalography (MEG) methods have demonstrated some facilitatory influence of musical expertise on pitch processing in music. For instance, the sensory and perceptual stages of pitch processing are typically improved in musicians compared to nonmusicians, as reflected by the enhancement of mid-latency (e.g., N19m–P30m, Schneider et al., 2002) and late electric and magnetic components (e.g., N1 and N1m, Pantev et al., 1998 and P2 and P2m, Shahin, Bosnyak, Trainor, & Roberts, 2003). Musical expertise also influences the automatic orienting responses to deviant sound features as reflected by the P3a component (e.g., Jongmsa, Desain, & Honing, 2004), as well as sound categorization and decision-related processes as reflected by the N2 component (e.g., Koelsch, Schröger, & Tervaniemi, 1999) and the late positivities of the P3 family (e.g., Besson & Faïta, 1995).

More controversial is the idea that musical expertise also influences pitch processing in speech (see Patel, 2008, for a review). Recent results have demonstrated more robust representations of pitch contour of Mandarin tones in the brainstem auditory responses (that reflect activity in the auditory nerve and subcortical structures) of musicians than of nonmusicians, even if none of the participants spoke Mandarin (Wong, Skoe, Russo, Dees, & Kraus, 2007). Moreover, the brainstem responses to both speech and music stimuli developed earlier and were larger in musicians than in nonmusicians (Musacchia, Strait, & Kraus, 2008; Musacchia, Sams, Skoe, & Kraus, 2007). Regarding cortical evoked potentials, enhanced preattentive processing (as reflected by the mismatch negativity, MMN) of pitch contour in musicians compared to nonmusicians has recently been shown by using iterated rippled noise with similar pitch contour than Tone 1 and Tone 2 in Mandarin Chinese (Chandrasekaran, Krishnan, & Gandour, 2009). The influence of musical expertise on attentive pitch processing in music and speech has also been examined by using parametric pitch manipulations of words/notes at the end of sentences and musical phrases (in children: Moreno et al., 2009; Magne, Schön, & Besson, 2006; Moreno & Besson, 2006; and in adults: Marques, Moreno, Castro, & Besson, 2007; Schön et al., 2004). Overall, results showed that musical expertise increased attentive pitch discrimination not only in music but also in speech, thereby showing positive music-to-language transfer effects. Results also revealed that the amplitude and latency of early (N1, N2/N3) and late (P3, late positivities) ERP components were larger and shorter in musicians compared to nonmusicians. Based on these previous findings, we hypothesized that tonal variations would elicit larger ERP effects with shorter onset latency in musicians than in nonmusicians. Moreover, and as mentioned above, contradictory results have been obtained regarding the influence of musical expertise on segmental/phonological processing. By analyzing the time course of the ERPs to segmental variations, we hoped to determine whether perceptual and/or cognitive processing would differ between musicians and nonmusicians.

## METHODS

### Participants

Twelve musicians and 12 nonmusicians gave their informed consent and were paid to participate in the experiment, which lasted for about 1 hr 30 min. Two participants (1 musician and 1 nonmusician) were not included in the analyses because of too many artifact-contaminated trials in their EEG recordings. Thus, the final groups comprised 11 musicians [mean age = 24 ( $SD = 2.6$ ) years, age range = 21–29 years, 7 women] and 11 nonmusicians [mean age = 23 ( $SD = 1.9$ ) years, age range = 21–28 years, 6 women]. All participants were right-handed, native speakers of French, with no experience of tone languages. They reported no hearing or neurological impairments. All participants had similar socioeconomic backgrounds and similar levels of education: All were students at the University and were either undergraduate or graduate students. None of the nonmusicians had any formal musical training (except for classic school education), and none of them played a musical instrument. Musicians played different instruments and started musical practice around the age of 7 ( $SD = 1.6$ ), with 16 years ( $SD = 3.8$ ) of musical training, on average, at the time of testing (see Table 1). None of them reported having absolute pitch.

### Materials

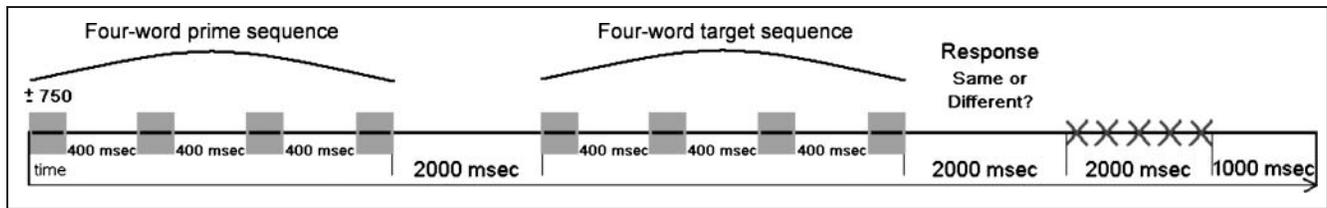
Two four-word sequences (prime and target sequences) were successively presented with a 2-sec time interval (see Figure 1). Mean word duration was 750 msec ( $SD = 216$  msec), the interword interval was 400 msec, and the mean duration of one sequence was 4312 msec ( $SD =$

390 msec). Sequences of words rather than single words were presented in order to increase the level of task difficulty. In order to avoid tone Sandhi, monosyllabic words were pronounced in isolation. Different words were used in the segmental session and in the tonal session, but within each session, the same 42 prime four-word sequences were presented both in the same and in the different conditions. A total of one hundred sixty-eight four-word sequences were presented within four blocks of 42 pairs each (tonal session: 2 blocks; segmental session: 2 blocks). Each block comprised 21 pairs of identical prime–target four-word sequences (same condition; 84 words) and 21 pairs of prime–target four-word sequences that differed by only one word (different condition; 21 words). Thus, a total of 105 different words were presented in each block. Because there were four blocks, a total of 420 spoken monosyllabic Mandarin Chinese words (sampling frequency = 44,100 Hz) were selected from Wenlin Software for Learning Chinese.

In the different condition, tonal or segmental variations were located on the final word for 30 target sequences, on the first word for four target sequences, on the second word for four target sequences, and on the third word for four target sequences (see Table 2). Only sequences with a variation on the last word of the target sequence were analyzed and the other sequences were used as distractors to prevent participants from paying attention to the last word only. Importantly, segmental variations in the different target sequences were always presented with the same tones as in the prime sequences so that only segmental information varied. Similarly for tonal variations, words in prime and target sequences were identical at the segmental level and

**Table 1.** Musical Background of Musicians

Musicians	Musical Training		
	Onset (Years)	Duration (Years)	Instruments
1	6	17	saxophone, guitar bass, accordion
2	8	17	double bass
3	6	11	violin
4	5	23	piano, flute
5	8	20	violin
6	5	14	piano, guitar bass
7	8	16	bassoon
8	10	15	clarinet
9	9	9	organ, harp
10	6	15	piano, flute, voice
11	7	17	saxophone, voice
Mean ( $SD$ )	7.1 (1.6)	15.8 (3.8)	



**Figure 1.** Illustration of the timing within a trial.

only tone varied. It is also important to note that tone information is typically available later than consonant information (Hallé, 1994) because most Mandarin Chinese words have a consonant–vowel structure and tones are primarily realized on vowels (Van Lancker & Fromkin, 1973). Thus, the onset of the acoustic differences was, on average, 143 msec later ( $SD = 80$ ) for tonal than for the segmental variations.

### Procedure

Sequences were presented through headphones in a pseudorandom order. Participants had to listen to the prime–target sequences in order to decide whether they were same or different and to press a response key as quickly and accurately as possible to give their response. The side (right or left hand) of the response was counter-balanced across participants. In the first two blocks, participants were asked to focus attention at the tonal level, and in the other two blocks, to focus attention at the segmental level. Furthermore, half of the participants started with the tonal session and the other half with the segmental session. Participants were informed that if there would be a variation, there would be only one in each target sequence and that it could occur on any of the four words of the target sequence.

Before the experimental phase, participants were familiarized with the “same–different” task and with the segmental and tonal variations on the Mandarin Chinese words. During the familiarization phase, participants were also trained to blink during the interstimulus interval in order to reduce the number of ocular artifacts in the experimental phase. The familiarization phase included four same pairs and four different pairs with tonal or segmental variations on Word 1 (1 sequence), on Word 2 (1 sequence), and on final word (2 sequences). Different words were used in the familiarization phase and in the experimental phase.

### ERP Recordings

Electroencephalogram (EEG) was continuously recorded from 32 BioSemi pin-type active electrodes (Amsterdam University), mounted on an elastic head cap, and located at standard left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas (International 10/20 System sites: Fz, Cz, Pz, Oz, Fp1, Fp2, AF3,

AF4, F3, F4, C3, C4, P3, P4, P7, P8, O1, O2, F7, F8, T7, T8, FC5, FC1, FC2, FC6, CP5, CP1, CP2, CP6, PO3, PO4). Moreover, to detect horizontal eye movements and blinks, the EOG was recorded from flat-type active electrodes placed 1 cm to the left and right of the external canthi, and from an electrode beneath the right eye. Two additional electrodes were placed on the left and right mastoids. EEG was recorded at a sampling rate of 512 Hz using BioSemi amplifier. The EEG was re-referenced off-line to the algebraic average of the left and right mastoids and filtered with a band pass of 0.1–40 Hz. Impedances of the electrodes never exceeded 5 k $\Omega$ . Data were segmented in single trials of 2200 msec starting 200 msec before the onset of the final word, and were analyzed using the Brain Vision Analyzer software (Brain Products, Munich). Trials with errors and trials containing ocular artifacts, movement artifacts, or amplifier saturation between 0 and 1000 msec were excluded from the averaged ERP waveforms.

### Data Analyses

A three-way analysis of variance (ANOVA), including expertise (musicians vs. nonmusicians) as a between-subject factor and session (segmental vs. tonal) and condition (same vs. different) as within-subject factors, was conducted on the behavioral data (error rates and RTs).

ERPs data were analyzed by computing the mean ERP amplitude relative to a 200-msec baseline in the 0–150, 150–250, 250–350, 350–600, 600–800, and 800–1000 msec latency bands as determined from previous results in the literature and from visual inspection of the waveforms. General ANOVAs were conducted separately for both midline and lateral electrodes and included expertise (musicians vs. nonmusicians) as a between-subject factor and session (segmental vs. tonal) and condition (same vs. different) as within-subject factors, together with electrodes (4 levels: Fz, Cz, Pz, and Oz) for midline electrodes. For lateral electrodes, the factors hemisphere (2 levels: left, right), regions of interest [ROI, 3 levels: fronto-central (F3, FC1, FC5 and F4, FC2, FC6), central (C3, CP1, CP5 and C4, CP2, CP6), and parietal (P3, P7, PO3 and P4, P8, PO4)] as well as electrodes (18 levels) were also included as within-subject factors. When the effects of session (segmental vs. tonal) were found to interact with the effects of expertise and/or condition, separate ANOVAs were conducted for the segmental and for the tonal sessions

**Table 2.** Tonal and Segmental Contrasts Used in the Experiment in the “Different” Condition

<i>Tonal Variations Used (Position in the Sequence)</i>		<i>Segmental Variations Used (Position in the Sequence)</i>	
<i>Prime (Word 4)</i>	<i>Target (Word 4)</i>	<i>Prime (Word 4)</i>	<i>Target (Word 4)</i>
fēi	féi	dié	tié
àng	áng	dái	zái
dóu	dōu	guō	guā
chè	chě	bán	zán
pǎng	pāng	suān	kuān
dūi	dùi	kùn	dùn
gǎi	gài	hà	dà
jiān	jiǎn	shǎng	chuǎng
yíng	yǐng	bà	bò
fán	fān	yàn	kèn
qìng	qíng	qué	lüé
dǔn	dún	zéi	léi
āi	ài	mǎi	dǎi
wǎn	wān	qiǔ	xiǔ
pìe	pǐe	zhà	shà
fèi	fěi	sōng	zhōng
zāo	záo	kuò	suò
píng	pīng	zhè	kè
kàn	kǎn	pài	pèi
zāi	zǎi	cǎi	zhǎi
yōu	yóu	chèn	yuàn
měi	mèi	xié	jié
kuān	kuàn	cǔn	lǔn
yǎ	yā	fǒu	zǒu
fò	fō	zōu	fōu
diú	diǔ	mīng	jīng
juě	jué	qún	kún
wéi	wěi	shāng	zhuāng
féng	fēng	wá	guá
qiǎn	qín	xiàn	tiàn
<i>Prime (Word 1)</i>	<i>Target (Word 1)</i>	<i>Prime (Word 1)</i>	<i>Target (Word 1)</i>
mù	mǔ	méi	léi
gě	gè	zhōu	chōu
zuǐ	zuī	tiān	diān
ròu	róu	zhǔ	zhǐ

**Table 2.** (continued)

<i>Tonal Variations Used (Position in the Sequence)</i>		<i>Segmental Variations Used (Position in the Sequence)</i>	
<i>Prime (Word 2)</i>	<i>Target (Word 2)</i>	<i>Prime (Word 2)</i>	<i>Target (Word 2)</i>
yū	yǔ	jīn	qīn
qiǎn	qián	qiǒng	yǒng
cū	cù	kú	tú
lǎ	lā	yūe	jūe
<i>Prime (Word 3)</i>	<i>Target (Word 3)</i>	<i>Prime (Word 3)</i>	<i>Target (Word 3)</i>
xǐng	xíng	hái	tái
suì	suī	xué	yué
bíng	bìng	shéng	céng
tú	tǔ	dàng	tàng

in latency bands adjusted from visual inspection to best capture the components elicited in each session. All  $p$  values were adjusted with Greenhouse–Geisser epsilon correction for nonsphericity. Tukey tests were used for post hoc comparisons.

## RESULTS

### Behavioral Data

Results showed lower error rates for musicians than for nonmusicians in both the tonal and segmental sessions [see Figure 2, top; main effect of expertise: musicians = 15.7% vs. nonmusicians = 22.7%;  $F(1, 20) = 4.87, p < .05$ , with no Expertise  $\times$  Session interaction:  $F(1, 20) = 1.38, p = .25$ ]. Moreover, participants made fewer errors in the segmental session (16.25%) than in the tonal session [22.25%; main effect of session:  $F(1, 20) = 16.54, p < .001$ ]. No other effect reached significance.

Results also showed faster RTs in the segmental session (1339 msec) than in the tonal session (1486 msec) in both groups [see Figure 2, bottom; main effect of session:  $F(1, 20) = 27.34, p < .001$ , with no Expertise  $\times$  Session interaction:  $F(1, 20) = 0.91, p = .35$ ]. No other effect reached significance.

### ERP Data

As can be seen in Figures 3 and 4, different words elicited larger early negative components (N1, N2/N3) and late positive components (P3a, P3b) than same words in both the segmental and tonal sessions, but with some interesting differences between the two sessions. Results of the general ANOVAs (see Table 3) revealed that the effect of session (tonal vs. segmental) interacted with the effects of the other factors of interest (i.e., specifically expertise and condition) in all latency bands considered for anal-

yses (except in the 0–150 msec latency band) at lateral, at midline, or at both midline and lateral electrodes. Consequently, separate ANOVAs were conducted for the segmental and tonal sessions (see Table 4).

In the segmental session (see Figure 3 and Table 4, top), no significant differences were found prior to 150 msec poststimulus onset.

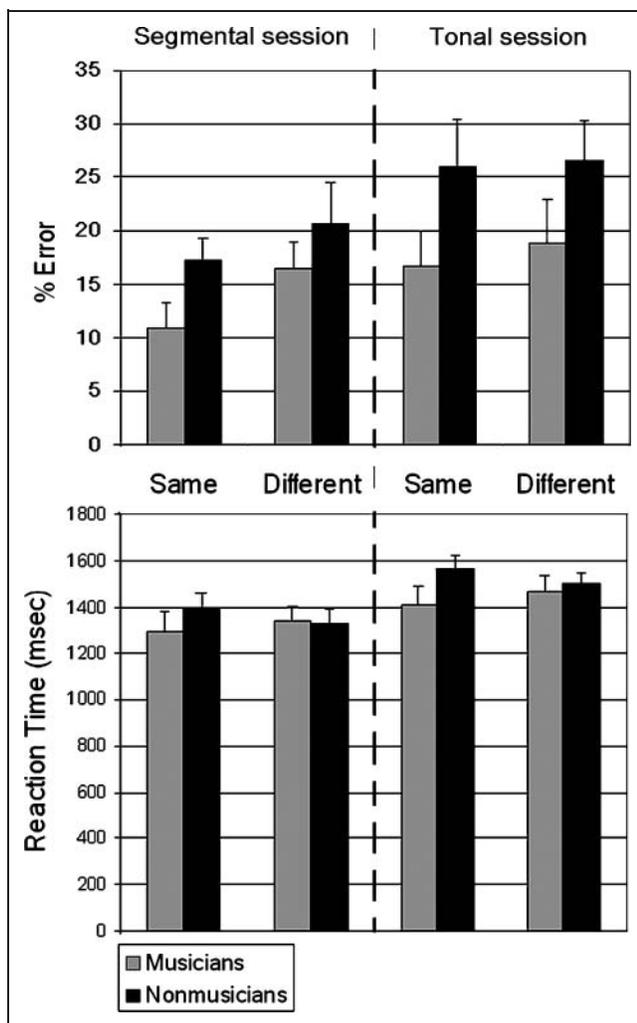
In the 150–250 msec range and in both groups, different words elicited larger N1 components at both midline ( $-1.35 \mu\text{V}$ ) and lateral electrodes ( $-1.56 \mu\text{V}$ ) than same words ( $0.1$  and  $-0.61 \mu\text{V}$ , respectively) over centro-parietal sites [Condition (C)  $\times$  Electrodes (Elec.),  $p = .03$  and C  $\times$  ROI,  $p = .02$ ].

In the 250–350 msec range and in both groups, different words ( $1.68 \mu\text{V}$ ) elicited larger P3a components than same words ( $0.63 \mu\text{V}$ ) at fronto-central lateral electrodes (C  $\times$  ROI  $\times$  Elec.,  $p = .01$ ).

In the 350–450 msec and 450–800 msec ranges, different words elicited larger P3b components ( $4.10 \mu\text{V}$  at midline and  $3.71 \mu\text{V}$  at lateral electrodes) than same words ( $2.11 \mu\text{V}$  at midline C,  $p < .01$  and  $1.76 \mu\text{V}$  at lateral electrodes: C,  $p < .01$ ) all over the scalp but with maximal amplitude at bilateral centro-parietal sites (C  $\times$  ROI,  $p = .03$  in the 350–450 msec range and C  $\times$  ROI,  $p < .005$  in the 450–800 msec range). Moreover, between 450 and 800 msec, the P3b effect (i.e., different – same words) was larger for musicians ( $2.73 \mu\text{V}$ ) than for nonmusicians ( $1.91 \mu\text{V}$ ) over bilateral parietal sites (Exp  $\times$  C  $\times$  R,  $p = .04$ ) with a bilateral scalp distribution in musicians and a left distribution in nonmusicians (Exp  $\times$  C  $\times$  H  $\times$  E,  $p = .05$ ; see P3 and P4 electrodes in Figure 3).

In the 800–1000 msec range, the P3b effect extended over bilateral parietal sites in both groups (C  $\times$  ROI,  $p = .02$ ).

In the tonal session (see Figure 4 and Table 4, bottom), no significant differences were found prior to 250 msec poststimulus onset. In contrast to the segmental session,



**Figure 2.** Percentage of errors (top) and reaction times (bottom) for musicians (gray) and nonmusicians (black) in the segmental session and in the tonal session and in the two experimental conditions (same and different with error bars).

the amplitude of the N1 component was not significantly different for same and different words.

In the 250–350 msec range and in musicians only, different words ( $-0.67 \mu\text{V}$ ) elicited larger N2/N3 components than same words ( $1.08 \mu\text{V}$ ) over bilateral parietal sites ( $\text{Exp} \times \text{C} \times \text{ROI} \times \text{E}, p = .04$ ).

Two different effects then developed in the 350–600 msec latency band. In the 350–450 msec range and in both groups, the N2/N3 component was larger for different ( $-1.0 \mu\text{V}$ ) than for same words ( $0.41 \mu\text{V}$ ) at Pz and over bilateral parietal sites (midline:  $\text{C} \times \text{E}, p = .03$ ; lateral electrodes:  $\text{C} \times \text{ROI}, p = .02$ ).

In the 450–600 msec range and in both groups, different words elicited larger P3a components at Fz ( $1.18 \mu\text{V}$ ) and over all lateral electrodes ( $1.35 \mu\text{V}$ ) than same words ( $-1.11$  and  $-0.08 \mu\text{V}$ , respectively; midline:  $\text{C} \times \text{E}, p = .04$  and lateral electrodes:  $\text{C}, p = .03$ ).

In the 600–800 msec range and in musicians only, different words ( $5.39 \mu\text{V}$  at midline and  $4.68 \mu\text{V}$  at lateral

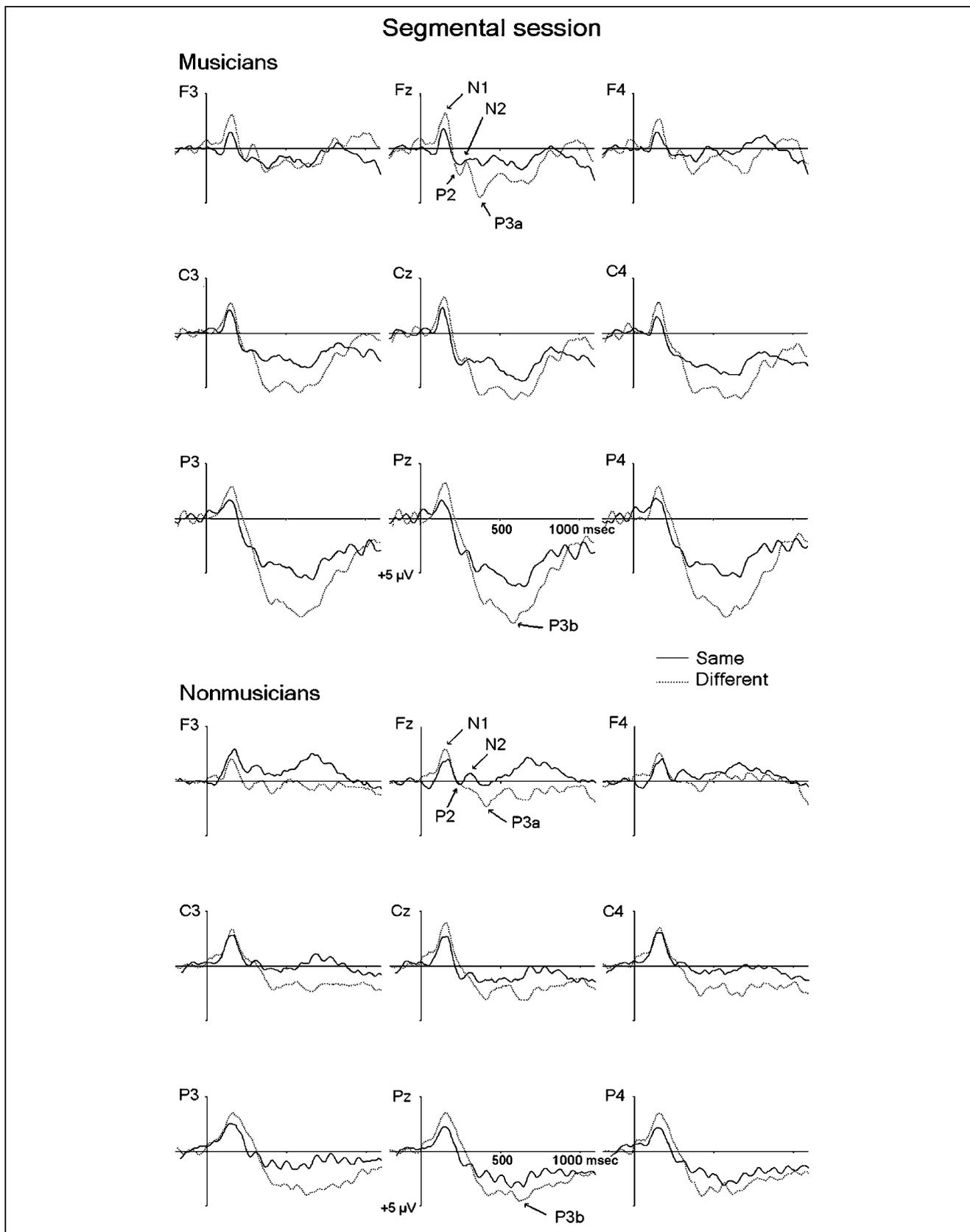
electrodes) elicited larger P3b components than same words ( $1.77 \mu\text{V}$  at midline and  $1.06 \mu\text{V}$  at lateral electrodes) over all scalp sites (cf.  $\text{Exp} \times \text{C}, p < .05$  at both midline and lateral electrodes).

Finally, in the 800–1000 msec range and in both groups, different words ( $3.74 \mu\text{V}$  at midline and  $3.40 \mu\text{V}$  at lateral electrodes) elicited larger P3b components than same words ( $1.02 \mu\text{V}$  at midline and  $0.67 \mu\text{V}$  at lateral electrodes;  $\text{C}, p < .001$  at both midline and lateral electrodes). Nevertheless, the P3b effect (i.e., *different* – *same words*) was significantly larger for musicians (*different* – *same* =  $3.96 \mu\text{V}$  at midline and  $3.85 \mu\text{V}$  at lateral electrodes) than for nonmusicians (*different* – *same* =  $1.49 \mu\text{V}$  at midline and  $1.59 \mu\text{V}$  lateral electrodes;  $\text{Exp} \times \text{C}, p = .02$  at both midline and lateral electrodes) and showed a widespread scalp distribution in musicians and a left centro-parietal distribution in nonmusicians ( $\text{Exp} \times \text{C} \times \text{H} \times \text{E}, p = .03$ ).

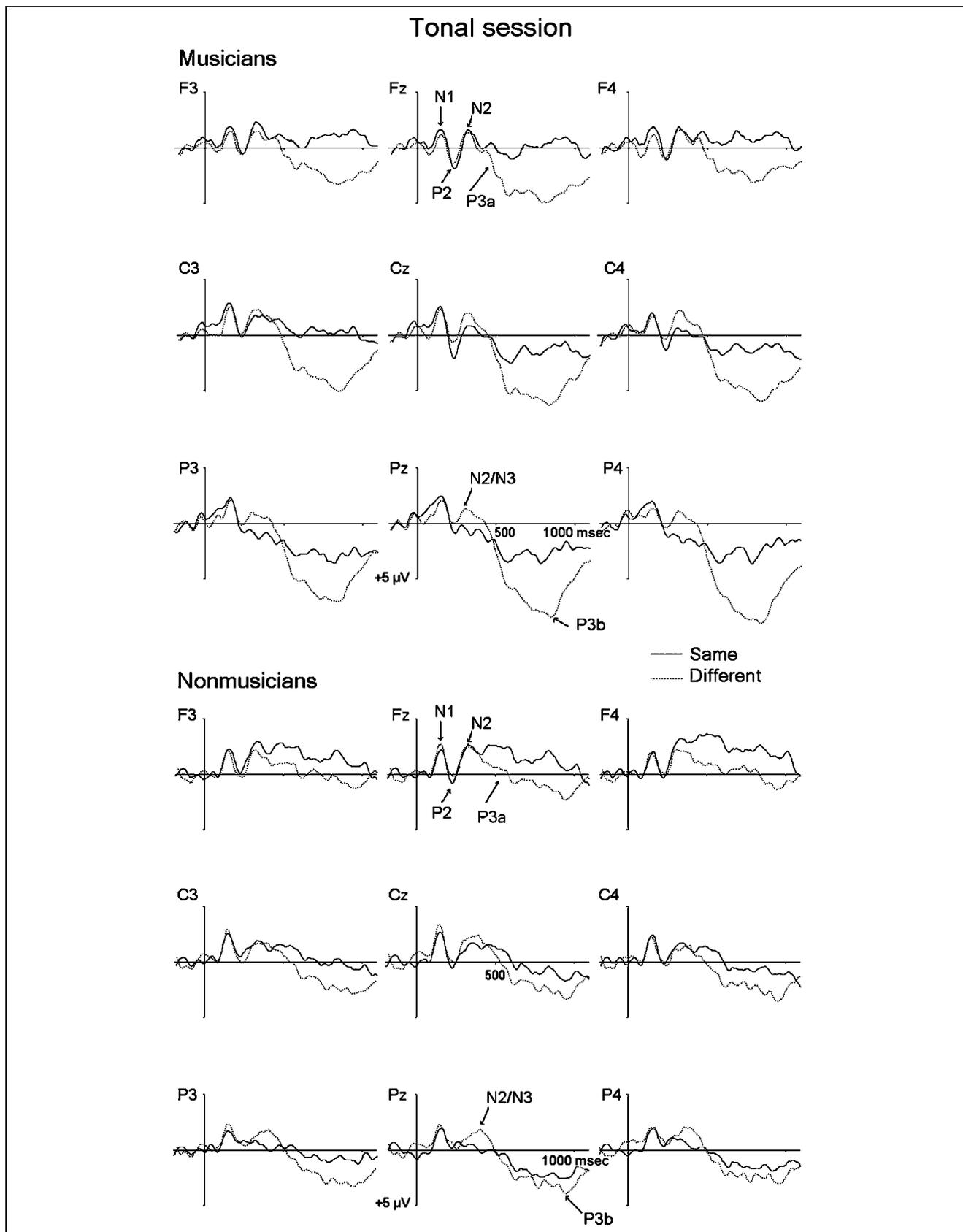
## DISCUSSION

In line with previous behavioral data, musicians showed higher discrimination of lexical tones in Mandarin Chinese than nonmusicians even if none of the participants were familiar with the language (Delogu et al., 2006, 2010; Lee & Hung, 2008; Alexander et al., 2005; Gottfried et al., 2004; Gottfried & Riester, 2000). Thus, as hypothesized, increased sensitivity to pitch acquired through years of musical practice increased lexical tone discrimination in a language in which tonal contrasts are linguistically relevant. At issue was whether musical expertise would also facilitate the processing of segmental variations. Interestingly, and in line with results showing a positive influence of musical skills on phonological processing (Jones et al., 2009; Moreno et al., 2009; Slevc & Miyake, 2006; Anvari et al., 2002; Lamb & Gregory, 1993), musicians also showed higher discrimination of segmental variations in Mandarin Chinese than nonmusicians. Taken together, these results provide further evidence of music-to-language transfer effects. However, these data do not allow us to specify which stages of processing (e.g., perceptual, cognitive or both) are influenced by musical expertise. Moreover, because the error rate was lower for musicians than for nonmusicians in all conditions, the influence of a general attentional factor cannot be ruled out. The electrophysiological data (see below) are useful in shedding light on these aspects.

Finally, and in line with earlier findings by Cutler and Chen (1997), participants made less errors in discriminating segmental than tonal variations. In French, although lexical pitch variations do not change the meaning of words, consonant or vowel variations do signal different words. Moreover, some segmental contrasts are similar in Mandarin and in French. Thus, it may be that French participants were more familiar with segmental than with tonal variations. In line with this interpretation, participants were also faster at discriminating segmental than tonal variations so that there was no speed–accuracy tradeoff. However,



**Figure 3.** ERPs in the segmental session time-locked to the onset of same (solid line) or different final words (dotted line) for musicians (top) and nonmusicians (bottom). In this figure, as in the following ones, the amplitude ( $\mu\text{V}$ ) is plotted on the ordinate (positive down) and the time (msec) is on the abscissa. ERPs from nine representative electrodes are presented.



**Figure 4.** ERPs in the tonal session time-locked to the onset of same (solid line) or different final words (dotted line) for musicians (top) and nonmusicians (bottom).

**Table 3.** Results of Statistical Analyses in the Different Latency Bands of Interest

Latency Bands (msec)	Electrodes	Factors	F	p
0–150	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	<i>ns</i>	<i>ns</i>	<i>ns</i>
150–250	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	S × C × R	$F(2, 40) = 3.5$	$p < .05$
250–350	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	Exp × S × C × H × R	$F(2, 40) = 6.20$	$p < .01$
350–600	Midline	C	$F(1, 20) = 7.69$	$p < .05$
		S × C × E	$F(3, 60) = 3.21$	$p < .05$
	Lateral	C	$F(1, 20) = 9.57$	$p < .01$
		S × C × R	$F(2, 40) = 3.21$	$p < .001$
600–800	Midline	C	$F(1, 20) = 26.7$	$p < .001$
		Exp × S × C	$F(1, 20) = 5.22$	$p < .05$
	Lateral	C	$F(1, 20) = 30.31$	$p < .001$
		Exp × C × R	$F(2, 40) = 4.41$	$p < .05$
800–1000	Midline	C	$F(1, 20) = 14.44$	$p < .01$
		Exp × S × C	$F(1, 20) = 6.44$	$p < .05$
	Lateral	C	$F(1, 20) = 15.79$	$p < .001$
		Exp × S × C	$F(1, 20) = 6.53$	$p < .05$

Exp = Expertise; S = Session; C = Condition; H = Hemisphere; R = ROI; E = Electrode; *ns* = nonsignificant.

segmental variations did also onset earlier, on the first phoneme, than tonal variations that occurred on the second phoneme. These differences in the onset of acoustic variations may also possibly account for the RT differences between the two sessions. Further experiments will aim at homogenizing the onset times of segmental and tonal variations to disentangle these effects.

### Electrophysiological Data

Analysis of the ERP components, N1 and N2/N3 components, as well as P3a and P3b components did provide precise information on the time course of the influence of musical expertise on the processing of tonal and segmental variations and revealed some interesting similarities and differences.

Considering tonal variations first, different words elicited larger N2/N3 components than same words over bilateral parietal sites. As can be seen in Figure 5, this difference started 100 msec earlier, on average, in musicians than in nonmusicians. Previous results have shown that the N2/N3 component is associated with stimulus discrimination and that its amplitude is larger and/or its latency longer for stimuli that are difficult to discriminate (e.g., Moreno et al., 2009; Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Ritter,

Simson, Vaughan, & Friedman, 1979; Simson, Vaughan, & Ritter, 1977). Based on this account, lexical pitch perception and discrimination, as reflected by the N2/N3 component, was faster in musicians than in nonmusicians, possibly because their increased sensitivity to pitch helped them discriminate different from same tonal words.

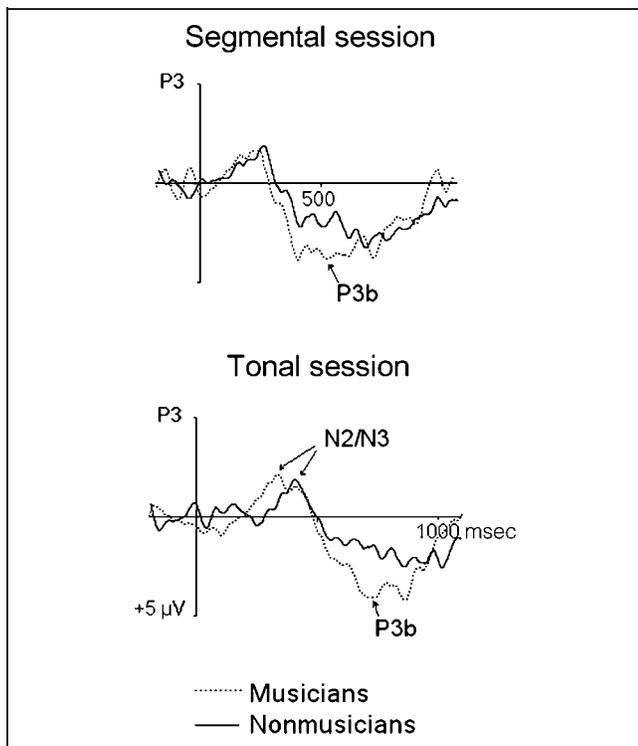
The N2/N3 component was followed by a fronto-central P3a component that was larger for different than for same words with no differences between musicians and nonmusicians. The P3a component is typically taken to reflect the automatic switching of attention toward unexpected and surprising events (e.g., Escera, Alho, Schröger, & Winkler, 2000; Courchesne, Hillyard, & Galambos, 1975; Squires, Squires, & Hillyard, 1975). Thus, the lexical pitch variations used here seemed to automatically attract musicians and nonmusicians' attention.

Finally, different words also elicited larger P3b components than same words and the latency of this effect was 200 msec shorter, on average, for musicians than for nonmusicians. Due to the continuous nature of ERPs recordings, the delayed P3b component in nonmusicians is probably tightly linked to the delayed N2/N3 components reported above in this group. Numerous results in the literature have shown that P3b is associated with categorization and decision processes (Picton, 1992; Duncan,

**Table 4.** Results of Statistical Analysis in the Segmental Session (Top) and in the Tonal Session (Bottom) in the Different Latency Bands of Interest

<i>Latency Bands (msec)</i>	<i>Electrodes</i>	<i>Factors</i>	<i>F</i>	<i>p</i>
<i>Segmental Session</i>				
0–150	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	<i>ns</i>	<i>ns</i>	<i>ns</i>
150–250 (N1)	Midline	C × E	$F(3, 60) = 3.55$	$p = .03$
	Lateral	C × R	$F(2, 40) = 5.26$	$p = .02$
250–350 (P3a)	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	C × R × E	$F(4, 80) = 3.92$	$p = .01$
350–450 (P3b)	Midline	C	$F(1, 20) = 13.77$	$p = .001$
	Lateral	C	$F(1, 20) = 11.96$	$p = .002$
		C × R	$F(2, 40) = 4.43$	$p = .03$
450–800 (P3b)	Midline	C	$F(1, 20) = 9.65$	$p = .005$
	Lateral	C	$F(1, 20) = 11.56$	$p = .003$
		C × R	$F(2, 40) = 8.33$	$p = .004$
		Exp × C × R	$F(2, 40) = 3.40$	$p = .04$
		Exp × C × H × E	$F(2, 40) = 3.09$	$p = .05$
800–1000 (P3b)	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	C × R	$F(1, 20) = 4.64$	$p = .02$
<i>Tonal Session</i>				
0–150	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	<i>ns</i>	<i>ns</i>	<i>ns</i>
150–250 (N1–P2)	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	<i>ns</i>	<i>ns</i>	<i>ns</i>
250–350 (N2/N3)	Midline	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Lateral	Exp × C × R × E	$F(4, 80) = 2.55$	$p = .04$
350–450 (N2/N3)	Midline	C × E	$F(3, 60) = 4.15$	$p = .03$
	Lateral	C × R	$F(2, 40) = 5.16$	$p = .02$
450–600 (P3a)	Midline	C × E	$F(3, 60) = 3.83$	$p = .04$
	Lateral	C	$F(1, 20) = 5.75$	$p = .03$
600–800 (P3b)	Midline	Exp × C	$F(1, 20) = 7.34$	$p = .01$
	Lateral	Exp × C	$F(1, 20) = 5.50$	$p = .03$
800–1000 (P3b)	Midline	C	$F(1, 20) = 29.82$	$p < .001$
		Exp × C	$F(1, 20) = 6.18$	$p = .02$
	Lateral	C	$F(1, 20) = 35.45$	$p < .001$
		Exp × C	$F(1, 20) = 6.07$	$p = .02$
		Exp × C × H × E	$F(2, 40) = 4.00$	$p = .03$

Exp = Expertise; C = Condition; H = Hemisphere; R = ROI; E = Electrode; *ns* = nonsignificant.



**Figure 5.** Difference waves (different minus same words) for the segmental (top) and tonal sessions (bottom) in musicians (dashed line) and nonmusicians (solid line).

Johnson & Donchin, 1977; see Donchin & Coles, 1988, for a review). Moreover, P3b amplitude and latency are known to be modulated by the level of task difficulty and by the level of confidence in the decision required by the task (Besson & Faïta, 1995; Parasurman, Richer, & Beatty, 1982; Squires et al., 1975; Squires, Hillyard, & Lindsay, 1973; Hillyard, Squires, Bauer, & Lindsay, 1971). In line with these previous findings and with the behavioral data, the present results showed that musicians categorized tonal variations more easily (shorter P3b latency) and were more confident in their decision (larger P3b amplitude) than nonmusicians. Finally, although the distribution of the P3b was widespread across scalp sites in musicians, it was more localized over left parietal sites in nonmusicians (see Figure 4). Recent results of an experiment with musicians that were native speakers of Mandarin showed left-lateralized late positivities to linguistic pitch variations in Mandarin Chinese and right-lateralized late positivities to music pitch variations (Nan, Friederici, Shu, & Luo, 2009). Thus, it may be that French nonmusicians relied more on segmental cues with greater involvement of the left hemisphere (Poldrack et al., 1999, 2001; Rugg, 1984), whereas French musicians relied both on left-lateralized phonological processing and on right-lateralized melodic and tone contour processing (e.g., Zatorre & Gandour, 2008; Zatorre & Krumhansl, 2002).

Turning to segmental variations, results showed that the N1 component that reflects the perceptual stages of information processing (Eggermont & Ponton, 2002;

Rugg & Coles, 1995) was larger to different than same words in both groups with no latency differences. Thus, both musicians and nonmusicians were equally sensitive to the segmental variations so that musical expertise did not seem to influence segmental perceptual processing. However, previous research has also shown that the N1 amplitude is modulated by attention (Hillyard, Hink, Schwent, & Picton, 1973). Thus, different segmental words may require more attention than same words. Although both interpretations are not mutually exclusive, the important point is that this N1 effect was similar for both musicians and nonmusicians.

As in the tonal session, segmental different words elicited larger fronto-central P3a components than same words that did not differ between musicians and nonmusicians, which is taken to show that the automatic orienting response (e.g., Courchesne et al., 1975) to segmental variations was not influenced by musical expertise.

By contrast, the amplitude of the P3b component was larger for different than for same words and for musicians than for nonmusicians (in the 450–800 msec range) with no latency differences. Based on the discussion above for tonal variations, this last result indicates that both musicians and nonmusicians were equally sensitive to segmental variations (no difference in P3b latency). However, in line with the behavioral data showing lower error rates for musicians than for nonmusicians, the larger P3b amplitude in musicians suggests that they were more confident than nonmusicians in their categorization of the words as same or different. Thus, cognitive rather than perceptual factors seem to explain the differences between musicians and nonmusicians to segmental variations. Finally, and similar to the tonal session, the scalp distribution of the P3b was parietal and left-lateralized in nonmusicians (see P3 and P4 on Figure 3). As argued above, this somewhat atypical left distribution of the P3b may be directly linked to the processing of segmental cues (Poldrack et al., 1999, 2001; Rugg, 1984).

It is important to note that the present results do not reflect a general effect of attention (i.e., musicians have better abilities to focus attention than nonmusicians) because some stages of lexical pitch processing and of segmental processing were influenced by musical expertise, whereas others were not. In the segmental session, musical expertise had no influence on perceptual processing of segmental variations (similar N1 components in both groups) but influenced the latency of the N2/N3 component that developed in the tonal session (shorter N2/N3 latency in musicians than in nonmusicians). Moreover, although the automatic orienting of attention toward segmental or tonal variations was similar in both groups (as reflected by the P3a component), musical expertise influenced the categorization and decision processes in both sessions (larger P3b amplitude and/or shorter latency in musicians than in nonmusicians).

Taken together, the present results extend previous findings in the literature (e.g., Chandrasekaran et al.,

2009; Moreno et al., 2009; Musacchia et al., 2008; Marques et al., 2007; Magne et al., 2006; Schön et al., 2004) that demonstrated music-to-language transfer effects and argue in favor of common pitch processing in music and speech. Along with results showing that some aspects of language and music processing, such as syntax and harmony for instance, may share some common processing resources (e.g., Fedorenko et al., 2009; Patel et al., 2008; Steinbeis & Koelsch, 2008; Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001), they argue against the idea of an encapsulated domain-specific language system (Pinker, 2005; Fodor, 1983). Finally, these results have interesting implications in the field of second language learning. We have recently shown that musical expertise also improves the discrimination of small and unusual increases in words' vowel duration (Marie, Magne, & Besson, 2010). Interestingly, in quantity languages such as Finnish, vowel duration has a contrastive function and can modify the meaning of words. By increasing the sensitivity to basic acoustic parameters such as pitch or duration, musical expertise may facilitate the building-up of higher-order representations (e.g., phonological) that are necessary to learn a foreign language.

## Conclusion

This study examined the influence of musical expertise on segmental and tone processing of words in a tone language unknown from the participants. We were able to demonstrate that several processes are involved in the same-different task, some of which are influenced by musical expertise (lexical pitch processing, categorization and decision processes) and some are not (segmental perceptual processing and automatic orienting of attention). These results add to previous findings in the literature by showing that musical expertise facilitates pitch perception in a tone language not only by refining the frequency processing network (Kraus, Skoe, Parbery-Clark, & Ashley, 2009) or the preattentive processing of sounds (Chandrasekaran et al., 2009) but also by improving word categorization and the level of confidence in the same-different decision. More generally, these results highlight the importance of following the time course of the different stages of information processing (some of which possibly going on in parallel) that allow naïve speakers to perceive different types of variations in a language that they do not understand. Taken together, these findings provide some scientific support to the folk idea that musicians can learn foreign languages more easily than nonmusicians. The next step could be to test if and how experience with a tone language modifies the perception of musical or nonspeech sounds.

## Acknowledgments

We thank Olivier Helleu for assistance in EEG recordings. This research was supported by a grant from the ANR-NEURO (024-01)

to M. Besson and was conducted at the "Institut de Neurosciences Cognitives de la Méditerranée."

Reprint requests should be sent to Céline Marie, Auditory Development Lab, Rm 329/D, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1, or via e-mail: mariec@mcmaster.ca.

## REFERENCES

- Alexander, J., Wong, P. C. M., & Bradlow, A. (2005). Lexical tone perception in musicians and nonmusicians. *Proceedings of Interspeech' Eurospeech—Ninth European Conference on Speech Communication and Technology*, Lisbon, Portugal.
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*, 111–130.
- Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1278–1296.
- Besson, M., & Schön, D. (in press). What remains of modularity? In P. Rebuschat, M. Rohrmeier, J. Hawkins, & I. Cross (Eds.), *Language and music as cognitive systems*. Oxford, UK: Oxford University Press.
- Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, *81*, 11–20.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, *108*, 1–9.
- Chao, Y. R. (1947). *Cantonese primer*. Cambridge, MA: Harvard University Press.
- Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*, 131–143.
- Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics*, *59*, 165–179.
- Delogu, F., Lampis, G., & Olivetti Belardinelli, M. (2006). Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, *7*, 203–207.
- Delogu, F., Lampis, G., & Olivetti Belardinelli, M. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, *22*, 46–61.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 355–425.
- Duncan-Johnson, C., & Donchin, E. (1977). On quantifying surprise, the variation of event-related potentials with subjective probability. *Psychophysiology*, *14*, 456–467.
- Eggermont, J., & Ponton, C. (2002). The neurophysiology of auditory perception: From single-units to evoked potentials. *Audiology and Neuro-Otology*, *7*, 71–99.
- Escera, C., Alho, K., Schröger, E., & Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neuro-Otology*, *5*, 151–166.
- Fedorenko, E., Patel, A. D., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, *37*, 1–9.

- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. A. (2000). *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*, 2593–2608.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, *12*, 207–222.
- Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *NeuroReport*, *9*, 2115–2119.
- Gottfried, T. L., & Riester, D. (2000). Relation of pitch glide perception and Mandarin tone identification. *Journal of the Acoustical Society of America*, *108*, 2604.
- Gottfried, T. L., Staby, A. M., & Ziemer, C. J. (2004). Musical experience and Mandarin tone discrimination and imitation. *Journal of the Acoustical Society of America*, *115*, 2545.
- Hallé, P. (1994). Evidence for tone-specific activity of the sternohyoid muscle in Modern Standard Chinese. *Language and Speech*, *37*, 103–124.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, *172*, 1357–1360.
- Jones, J. L., Lucker, J., Zalewski, C., Brewer, C., & Drayna, D. (2009). Phonological processing in adults with deficits in musical pitch recognition. *Journal of Communication Disorders*, *42*, 226–234.
- Jongsma, M., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. *Biological Psychology*, *66*, 129–152.
- Klein, D., Zatorre, R. J., Milner, B. A., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, *13*, 646–653.
- Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*, 1309–1313.
- Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre and timing: Implications for language and music. *Annals of the New York Academy of Sciences: Neurosciences and Music III*, *1169*, 543–557.
- Lamb, S. J., & Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Journal of Educational Psychology*, *13*, 13–27.
- Lee, C. H., & Hung, T. H. (2008). Identification of Mandarin tones by English speaking musicians and non-musicians. *Journal of the Acoustical Society of America*, *124*, 3235–3248.
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than non-musician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*, 199–211.
- Marie, C., Magne, C., & Besson, M. (2010). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*, 294–305.
- Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than non-musicians: Behavioural and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*, 1453–1463.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36–47.
- Moreno, S., & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology*, *43*, 287–291.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*, 712–723.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences, U.S.A.*, *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 nonmusicians. *Hearing Research*, *241*, 34–42.
- Nan, Y., Friederici, A. D., Shu, H., & Luo, Y. J. (2009). Dissociable pitch processing mechanisms in lexical and melodic contexts revealed by ERPs. *Brain Research*, *1263*, 104–113.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.
- Parasuraman, R., Richer, F., & Beatty, J. (1982). Detection and recognition: Concurrent processes in perception. *Perception & Psychophysics*, *31*, 1–12.
- Patel, A. D. (2008). *Music language and the brain*. Oxford, New York: Oxford University Press.
- Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*, 456–479.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pinker, S. (2005). So how does the mind work? *Mind and Language*, *20*, 1–24.
- Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M. M., et al. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI. *Journal of Cognitive Neuroscience*, *13*, 687–697.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, *10*, 15–35.
- Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, *94*, 67–78.
- Ritter, W., Simson, R., Vaughan, H. G., & Friedman, D. (1979). A brain event related to the making of sensory discrimination. *Science*, *203*, 1358–1361.
- Rugg, M. D. (1984). Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia*, *22*, 435–443.
- Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event related brain potentials and cognition* (pp. 27–39). Oxford, UK: Oxford University Press.
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*, 688–694.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music facilitates pitch processing in language. *Psychophysiology*, *41*, 341–349.
- Shahin, A., Bosnyak, D., Trainor, L., & Roberts, L. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, *23*, 5545–5552.

- Simson, R., Vaughan, H. G., & Ritter, W. (1977). The scalp topography of potentials in auditory and visual discrimination tasks. *Electroencephalography and Clinical Neurophysiology*, *42*, 528–535.
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second language proficiency: Does musical ability matter? *Psychological Science*, *17*, 675–681.
- Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973). Vertex potentials evoked during auditory signal detection: Relation to decision criteria. *Perception & Psychophysics*, *14*, 265–272.
- Squires, K. C., Squires, N. K., & Hillyard, S. A. (1975). Decision-related cortical potentials during an auditory signal detection task with cued observation intervals. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 268–279.
- Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, *18*, 1169–1178.
- Van Lancker, D., & Fromkin, V. A. (1973). Hemispheric specialization for pitch and “tone”: Evidence from Thai. *Journal of Phonetics*, *1*, 101–109.
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.
- Xu, Y. (2001). Sources of tonal variations in connected speech. *Journal of Chinese Linguistics Monograph Series*, *17*, 1–31.
- Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, *33*, 319–337.
- Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations of speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, *363*, 1087–1104.
- Zatorre, R. J., & Krumhansl, C. L. (2002). Mental models and musical minds. *Science*, *298*, 2138–2139.