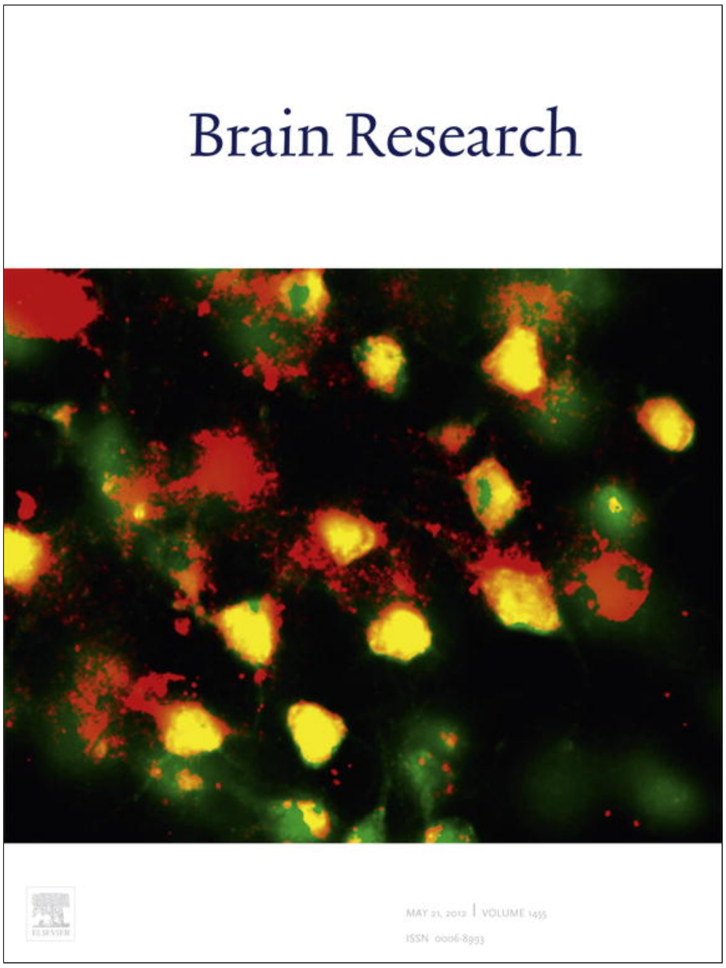
Brain Research
MAY 21, 2012 | VOLUME 1455
ISSN 0006-8993

BRAIN RESEARCH 1455 (2012) 75–89

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study

Chizuru Deguchi[a,c,*], *Magali Boureux*[a,d], *Michela Sarlo*[a], *Mireille Besson*[b], *Massimo Grassi*[a], *Daniele Schön*[b], *Lucia Colombo*[a]

[a]*Dipartimento di Psicologia Generale, University of Padova, Italy*
[b]*Institut de Neurosciences Cognitives de la Méditerranée, CNRS-Marseille, Marseille, France*
[c]*Dipartimento di Scienze della Cognizione e della Formazione, University of Trento, Italy*
[d]*Dipartimento di Filosofia, Pedagogia e Psicologia, University of Verona, Italy*

A R T I C L E I N F O

Article history:
Accepted 15 March 2012
Available online 23 March 2012

Keywords:
Pitch detection
Sentence prosody
Musical expertise
ERP
Jabberwocky

A B S T R A C T

Previous ERP studies have shown that musicians detect a pitch change in spoken sentences better than non-musicians in both native (French, Schön et al., 2004) and unfamiliar (Portuguese, Marques et al., 2007) language. The aim of the present study was to further investigate differences between musicians and non-musicians in processing pitch changes in spoken sentences. To study the effects of familiarity of intonational contour and of the presence of meaningful context, behavioral and electrophysiological data from Italian musicians and non-musicians were compared in a pitch incongruity detection task using sentences in the native (Italian) and foreign (French) language and in jabberwocky (meaningless sentences formed by pseudowords). Moreover, to examine whether these differences depend on enhanced auditory sensitivity to pitch, the frequency discrimination threshold (FDT) for tones was obtained using a psychophysical procedure. Musicians were more accurate than non-musicians in detecting small pitch changes in all languages showing a smaller response bias, as well as much lower FDTs than non-musicians. The ERP data revealed shorter latencies of a late positivity over parietal sites in musicians than in non-musicians for weak and strong incongruities. Overall results confirmed musicians' advantage in detection of subtle pitch changes not only with tones but also with speech sentences in both native and unfamiliar languages. Such effect appears to emerge from more efficient pitch analysis trained by musical experience.

1. Introduction

A number of recent studies have examined whether musical experience shapes perceptual processing similarly in music and language. Perceptual abilities involved in the perception of the temporal structure (i.e., meter, rhythm) and of the fundamental frequency (f0, or pitch) of a sound sequence have been investigated. Pitch is a fundamental dimension in

* *Corresponding author at:* Dipartimento di Scienze della Cognizione e della Formazione, University of Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy. Fax: +39 046 480 8602.
E-mail address: chizuru.deguchi@unitn.it (C. Deguchi).

0006-8993/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2012.03.034

speech because it contributes to prosodic information present in spoken language at different levels (i.e., syllable, words and sentences) and in music to identify a melody, and to provide one of the main features of the specific aesthetic and emotional quality of a piece of music. Of main interest here is the perception of pitch contour, a critical aspect of both speech and music that defines the "melody" of a sequence of sounds.

Positive effects of musical expertise on tonal pitch processing have been shown in previous psychophysical experiments (Kishon-Rabin et al., 2001; Micheyl et al., 2006; Spiegel and Watson, 1984). Micheyl et al. (2006) found that thresholds for discriminating pure and complex harmonic tones were six times lower and with smaller variability for musicians compared to non-musicians. Interestingly, only from four to six hours of training in the psychoacoustic task was sufficient to halve difference in performance between groups: the training could improve performance of non-musicians, but did not provide further benefit to musicians. Using electrophysiological measures, results of many experiments have demonstrated enhanced pre-attentive and attentive processing of harmonic and musical sounds in musicians (e.g., Chandrasekaran and Kraus, 2010).

Recent studies have supplied evidence that musical expertise provides an advantage in differentiating pitch contours, not only in music but also in language, using Event-Related brain Potentials (ERPs). Thanks to its excellent temporal resolution, the ERP technique allows us to track neural responses to speech or musical materials, which typically develop in time. The benefit of musical expertise has been reflected in more accurate brainstem encoding of pitch direction for musicians than for non-musicians listening to Mandarin tones with which they were not familiar (e.g., Wong et al., 2007). At the cortical level, musicians showed increased MMN responses to non-speech stimuli simulating acoustic pitch shifts of Mandarin tones (e.g., Chandrasekaran et al., 2009), and Musacchia et al. (2008) have reported correlations between brainstem and early cortical ERPs. Recent results have also shown that the level of performance in same-different tasks on sequences of Mandarin monosyllabic words varying in tone or segmental phonology was higher in musicians than non-musicians, and this was reflected in larger and/or earlier N2 and P3 components of the ERPs (Marie et al., 2011).

Directly related to our present concerns, detection of changes in pitch height (f0) embedded in spoken sentences was compared between French musicians and non-musicians listening to their native language (Schön et al., 2004) or to a foreign language (Portuguese; Marques et al., 2007). The final word of each sentence could be intact ('congruous') or carry a small ('weak incongruity') or large increase in f0 ('strong incongruity'). In both studies, musicians were faster and made fewer errors than non-musicians in detecting small pitch changes. Moreover, electrophysiological results revealed that the positive component over parietal electrodes elicited by strong and weak f0 changes showed an earlier onset in musicians than in non-musicians. Overall, these data suggest that, due to extensive training with dynamic melodic pitch patterns, musicians' greater perceptual ability to perceive pitch changes in music generalizes to the perception of pitch contour in speech prosody, not only in the native language but also in a foreign language. However, some differences also emerge from the comparison of the two studies. For both musicians and non-musicians, reaction times were longer, and the positive ERP component did onset later, for Portuguese than for French sentences, thereby suggesting that discriminating pitch changes in a foreign language, particularly when they are subtle, increased task difficulty. This effect was only partially compensated by musical expertise.

To explain the differences of results between the native and foreign languages, Marques et al. (2007) emphasized the importance of semantic information. Being able to understand the French sentences made it possible for the French participants to anticipate when the last word of the sentence, carrying the pitch variation, was presented. However, assuming that speakers implicitly learn typical prosodic patterns in their native language, another possible interpretation is that being familiar with an intonational contour can possibly help in anticipating and detecting strong or subtle pitch changes in the native language. By contrast, anticipating such changes in a foreign language is more demanding, since the intonational contour is not familiar.

This alternative interpretation was addressed in a previous study (Colombo et al., 2011) in which Italian and French sentences, with congruous, weakly or strongly incongruous pitch changes on the final word, were presented to non-musician native speakers of Italian who had never formally studied French.[1] Importantly, the between-language comparison was made within-subjects, in order to avoid biases due to group differences. Moreover, to evaluate the relative contribution of semantic and prosodic factors jabberwocky sentences, that is, sentences formed by "legal" Italian nonwords with no meaning, but with intonational contour similar to the homologous Italian sentences, were also included. Results showed that familiarity with the intonational contour of sentences in the native language was helpful in detecting both strong and weak pitch changes (i.e., sensitivity to pitch changes was higher for jabberwocky than for French sentences). However, Italian sentences maintained an advantage compared to jabberwocky, thereby underlying that participants

[1] Although these two languages present some lexical similarities, Italian participants could not extract from French sentences but only sporadic lexical items that were insufficient for understanding the phrasal context or for guessing the final word. Such difficulty may be due to the fact that Italian and French listeners exploit different cues to segment speech into significant units (e.g., words). For example, position of word stress is free in Italian, and has a distinctive lexical value, while in French it is always on the last syllable. Accordingly, while French listeners use a metrical segmentation strategy based on the final syllable of the word, setting a boundary immediately after it (Banel and Bacri, 1994, 1997), Italian listeners use other types of cues, for example, distributional cues such as the correlation between sounds within- and between-syllables and frequency of occurrence of syllable (Caniparoli, 2001). Moreover, language-specific phenomena such as the "resyllabation" in French, for which the final consonant is pronounced in connection with the following vowel (e.g. "petit lit" is pronounced */pti/li/* while "petit ami" is pronounced */pti/ta/mi/*, not */ptit/a/mi/* nor */pti/a/mi/*), are considered to augment segmentation difficulty (Delattre, 1947, 1966).

usefully exploited the semantics of the sentence in order to carry out the task.

Based on these results, the first aim of the present experiment was to examine the influence of musical expertise by presenting the same materials and linguistic conditions to musicians and non-musicians who were all native Italian speakers and did not know French. Based on previous results, we expected Italian musicians to perform better than non-musicians both in Italian and in French. Of main interest here was to determine whether this advantage would extend to jabberwocky sentences. The comparison between Italian and jabberwocky, on one hand, and between jabberwocky and French, on the other hand, would provide information as to whether musicians can better exploit the meaningful context and/or familiar pitch contour, respectively, than non-musicians in the pitch detection task.

In order to evaluate participants' auditory abilities in discriminating pitch differences independently of any linguistic context, a frequency (f0) discrimination threshold (FDT; Micheyl et al., 2006) for complex tones was also measured using a psychoacoustic procedure with a two-alternative forced choice task (2AFC). Overall, we expected musicians to outperform non-musicians (Kishon-Rabin et al., 2001; Micheyl et al., 2006). Of more specific interest here was to determine whether discrimination threshold in the psychoacoustic task would correlate with behavioral performance (RTs and error rate) in the sentence pitch detection task and whether such correlational pattern would be similar or different for musicians and for non-musicians. Such results may throw light on the processes that are responsible for musicians' advantage. Finding strong correlations between the FDT for isolated pairs of harmonic tones and behavioral measures in the sentence pitch detection task would argue in favor of enhanced sensory ability in pitch perception in musicians compared to non-musicians being extended to a non-musical task.

However, the musicians' advantage in the prosody task may result not only from enhanced perceptual sensitivity, but also from more efficient cognitive post-perceptual decision processes to perform the task. While the FDT measure obtained by a forced choice task returns the net measure of sensory performance unaffected by response bias (Green and Swets, 1966), behavioral performance in the prosody task can be modulated by the listener's attitude and strategies employed in response decision. In order to specifically address this issue, a signal detection theory (SDT) analysis was carried out on the behavioral data, by calculating the relative size of d′, the index of perceptual sensitivity, and c, the criterion for response bias (Green and Swets, 1966).

Finally, based on previous results (Colombo et al., 2011), we examined early auditory processing and later cognitive processes in the sentence pitch detection task by analyzing the early negative component and following positivity elicited by sentence final words. As the strong incongruity was perceptually salient, it was expected to elicit a fronto-temporal perceptual N1 component, peaking around 150 ms after pitch change onset (Eggermont and Ponton, 2002; Rugg and Coles, 1995) followed by a parietal positive component (P3), peaking around 300–350 ms after pitch change onset and reflecting attention shift to perceptually salient and cognitively relevant pitch changes (Escera et al., 2000) or updating of working memory (Donchin and Coles, 1988). By contrast, the weak incongruity was perceptually less salient, and the fronto-temporal negativity and parietal positivity were expected to be smaller in amplitude and to develop with longer latency than to strongly incongruous endings (Duncan-Johnson and Donchin, 1977; Fujioka et al., 2006; Moreno et al., 2009; Picton, 1992). Finally, we expected these differences to be larger in musicians than in non-musicians, specifically in the most difficult conditions (jabberwocky and French sentences).

To summarize, the aims of the present study were to determine whether the sensory and cognitive abilities deployed by the participants in a sentence pitch detection task differ between musicians and non-musicians. To this end, we analyzed behavioral data, as well as early and late ERP components. We also compared musicians and non-musicians using a psychoacoustic measure of pitch discrimination acuity with complex tones (FDT) as a measure of auditory sensitivity to pitch free from any linguistic context. Finally, we used different language contexts to pinpoint the influence of the familiarity with sentence prosody and semantics.

2. Results

2.1. *Behavioral data*

ANOVAs were conducted on RTs for correct responses only and on error rates after outlier correction (RTs longer than 2500 ms, 1.3%), with Congruity of the target prosody with the sentence context (congruous, weakly incongruous, strongly incongruous) and Language (Italian, jabberwocky and French) as within-subject factors. Expertise (musicians vs. non-musicians) was a between-subjects factor.

Mean error rates and mean reaction times (RTs) for the two groups of participants are shown in Fig. 1.

The ANOVA on error rates showed significant main effects of Language [$F(2,52)=20.38$, MSE=.008, $p<.001$], Congruity [$F(2,52)=62.42$, MSE=.04, $p<.001$] and Expertise [$F(1,26)=7.41$, MSE=.029, $p<.05$]. The Congruity by Expertise [$F(2,52)=6.80$, MSE=.04, $p<.01$] and Congruity by Language interactions [$F(2,52)=9.62$, MSE=.007, $p<.001$] were also significant. Planned comparisons showed that the error rate was significantly lower for musicians than for non-musicians in the weakly incongruous condition (25% vs. 44%, $p<.001$) but did not differ between groups in the congruous and strongly incongruous conditions. Differences between languages were found only in the weakly incongruous condition: error rate was significantly higher for French (46%) than for jabberwocky (31%, $p<.01$) and for Italian (27%, $p<.001$) while the Italian–jabberwocky difference was not significant.

The ANOVA on RTs showed no main effect of Expertise, nor any interaction of Expertise with other factors. The main effects of Language [$F(2,52)=25.16$, MSE=19,349, $p<.001$] and Congruity [$F(2,52)=106.46$, MSE=36,568, $p<.001$] and the Language by Congruity interaction were significant [$F(4,104)=3.69$, MSE=9328, $p<.01$]. Planned comparisons showed that RTs were longer for French than for jabberwocky and for Italian in the congruous and weakly incongruous conditions and for French than for Italian in the strongly incongruous condition.

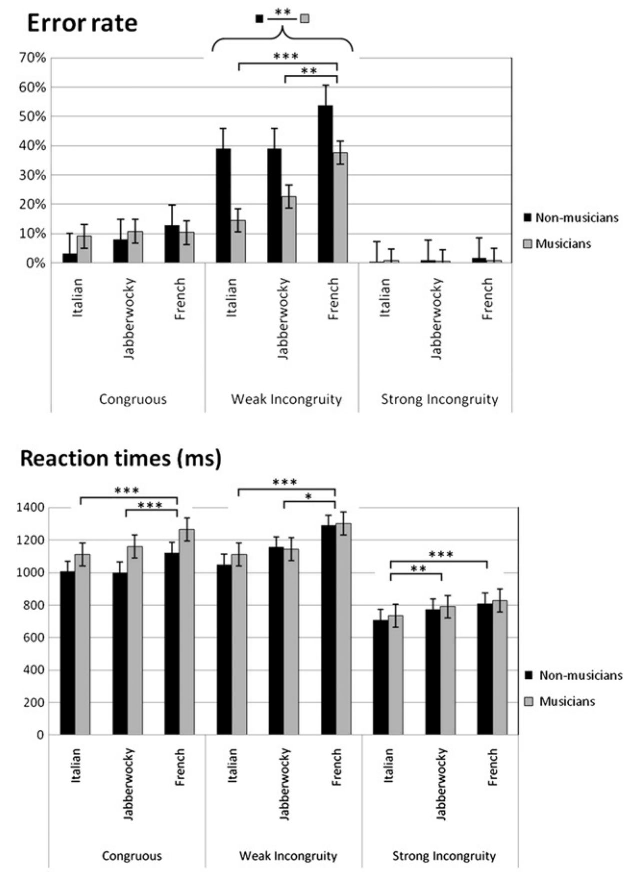

Fig. 1 – Mean error rates (upper panel) and mean reaction times (lower panel) for 14 non-musicians (black) and 14 musicians (gray) in the pitch detection task using spoken sentences in Italian, jabberwocky and French.

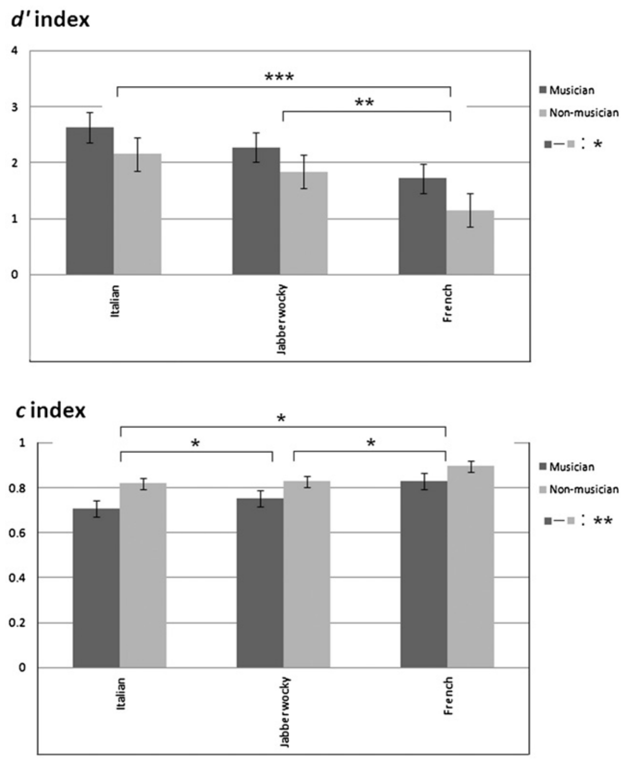

Fig. 2 – Mean values of d′ (upper panel) and c (lower panel) indexes for 14 non-musicians (black) and 14 musicians (gray) for weak incongruity in Italian, jabberwocky and French in the sentence pitch detection task.

To investigate the perceptual and response decision process in judging small pitch changes that are difficult to perceive, behavioral data for weak incongruity were further analyzed by considering two indexes from signal detection theory: d′ (reflecting perceptual sensitivity) and c (reflecting response criterion; Green and Swets, 1966).

In the analysis of d′ (see Fig. 2, upper panel), using ANOVA with Language (three levels) as within-subject factor and Expertise (two levels) as between-subject factor, the main effects of Language [$F(2,52)=18.42$, $MSE=.36$, $p<.001$] and Expertise [$F(1,26)=5.59$, $MSE=.92$, $p<.05$] were significant, with larger d′ for Italian (mean $d'=2.39$, $SD=0.13$) and for jabberwocky (mean $d'=2.05$, $SD=0.16$) than for French (mean $d'=1.43$, $SD=0.13$). Sensitivity was higher in musicians (mean $d'=2.21$, $SD=0.15$) than in non-musicians (mean $d'=1.71$, $SD=0.15$), and was higher in sentences with familiar semantics and prosody (Italian) or only with familiar prosody (jabberwocky) compared to French. Thus, expected differences in sensitivity showed up, when pitch changes were perceptually difficult to discriminate.

As a measure of response bias, we used c (Green and Swets, 1966; see Fig. 2, lower panel). In the ANOVA with Language (three levels) as within-subject factor and Expertise (two levels) as between-subject factor, the main effects of Language [$F(2,52)=21.68$, $MSE=.003$, $p<.001$] and Expertise [$F(1,26)=7.96$, $MSE=.02$, $p<.01$] were significant, with no interaction. Compared to musicians (mean $c=0.76$, $SD=0.021$), non-musicians (mean $c=0.85$, $SD=0.021$) showed a larger bias to give a negative response to weak pitch changes, thus evaluating them as congruous. In both groups there was a larger bias to give a negative response in French (mean $c=0.86$, $SD=0.017$), then in jabberwocky (mean $c=0.79$, $SD=0.018$) and then in Italian (mean $c=0.76$, $SD=0.017$). In sum, compared to musicians, non-musicians were less sensitive and had a larger bias to evaluate the weakly incongruous endings as congruous.[2]

2.2. *ERP results*

ANOVAs on mean ERP amplitudes were separately computed for each language because the time course of the components was very different for the three languages, in particular for Italian and French, and also because the comparison among languages was not the main aim of the present study (but see Colombo et al., 2011).

[2] Although the signal detection theory states that d′ and c are independent measures (Ingram, 1970; MacMillan, 1993; McNicol, 1972), this independence does not always find empirical support. For this reason, the comparative analysis using d′ and c results should be taken cautiously.

Results of ANOVAs for midline and lateral electrodes and for the three languages are reported in Table 1. Only the statistically significant effects and interactions involving Expertise or/and Congruity are reported here. The ERP grand-averages for Italian, jabberwocky and French are shown in Figs. 3, 4 and 5, respectively.

In the 150–250 ms range, the Congruity by ROI interaction (lateral analysis) was significant for all languages. Strongly incongruous endings elicited larger early negativity (the N1 component) than weakly incongruous and congruous endings over fronto-central and temporal regions (always $p<.001$). Moreover, the N1 component was larger over the right than the left hemisphere (Congruity by Hemisphere interaction).

In the 250–750 ms range, the P3 component was larger for strong incongruity than for weak incongruity and congruous endings and larger over parietal than fronto-central regions (Congruity by Electrode and Congruity by ROI interactions in the midline and lateral analyses, respectively).

Finally, in the 750–1000 ms range, the amplitude of the late positivity at parietal sites was larger for the weak incongruity than for both the congruous endings and strong incongruity, whereas the latter two conditions did not differ from each other (Congruity by Electrode and Congruity by ROI interactions in midline and lateral analyses, respectively).

In order to determine the onset latency of the congruity effects (i.e., when ERPs in the strongly and weakly incongruous conditions start to diverge from the ERPs in the congruous condition) in each of the three languages and for both musicians and non-musicians, the mean amplitude of the positive component was measured at Pz where it was largest (see Tables 2 and 3). Results revealed that for Italian, in the weak incongruity condition (see Table 2), the onset latency of the positivity was 300 ms shorter for musicians (400–450 ms) than for non-musicians (700–750 ms). For jabberwocky, in the strong incongruity condition (see Table 3), the onset latency was 100 ms shorter for musicians (200–250 ms) than for non-musicians (300–350 ms). No between-group difference was found for French.

Table 1 – Results of ANOVAs for Italian, jabberwocky and French sentences on mean amplitudes of the ERPs at lateral electrodes.

Language		Latency Band (ms)	Factors	df	F	p
Italian	Midline	150–250	(ns)			
		250–750	C	2, 40	40.15	***
		250–750	C×E[a]	4, 80	14.27	***
		750–1000	C	4, 80	12.13	***
		750–1000	C×E[b]	4, 80	6.09	**
	Lateral	150–250	C	2, 40	7.72	**
		150–250	C×H[c]	2, 40	6.57	**
		150–250	C×R[d]	4, 80	4.61	*
		250–750	C	2, 40	34.05	***
		250–750	C×R[e]	4, 80	23.51	***
		750–1000	C	2, 40	9.63	**
		750–1000	C×R[f]	4, 80	5.88	**
Jabberwocky	Midline	150–250	(ns)			
		250–750	C	2, 40	68.43	***
		250–750	C×E[a]	4, 80	8.99	***
		750–1000	C	4, 80	3.65	*
		750–1000	C×E[b]	4, 80	3.07	*
	Lateral	150–250	C	2, 40	5.03	*
		150–250	C×H[c]	2, 40	4.54	*
		150–250	C×R[d]	4, 80	5.59	*
		150–250	C×R×H	4, 80	4.4	*
		250–750	C	2, 40	59.07	***
		250–750	C×R[e]	4, 80	14.94	***
		750–1000	C×R[f]	4, 80	5.87	**
French	Midline	150–250	(ns)			
		250–750	C	2, 40	42.88	***
		250–750	C×E[a]	4, 80	24.43	***
		750–1000	(ns)[b]			
	Lateral	150–250	C	2, 40	5.63	**
		150–250	C×H[c]	2, 40	3.43	*
		150–250	C×R[d]	4, 80	4.33	*
		250–750	C	2, 40	45.15	***
		250–750	C×R[e]	4, 80	35.59	***
		750–1000	C×R[f]	4, 80	4.4	*

Notes to Table:

Exp: musical expertise (musicians vs. non-musicians);
C: Congruity (Congruous vs. Weak incongruity vs. Strong incongruity); R: ROI (region of interest); H: hemisphere;
E: electrodes.df: uncorrected degrees of freedom;
p: probability level after Greenhouse–Geisser correction of F-values.

[a] The mean amplitude at Pz was more positive for strong incongruity than for other conditions: Italian = 10.20 μV, 5.94 μV, 3.63 μV; Jabberwocky = 8.94 μV, 3.42 μV, 2.68 μV; French = 8.86 μV, 1.96 μV, 1.98 μV, for strong incongruity, weak incongruity and congruous endings, respectively.

[b] The mean amplitude at Pz was more positive for weak incongruity than for other conditions in Italian and Jabberwocky: Italian = 5.49 μV, 9.49 μV, 4.71 μV; Jabberwocky = 4.83 μV, 7.27 μV, 4.65 μV; French = 5.63 μV, 7.73 μV, 5.53 μV, for strong incongruity, weak incongruity and congruous endings, respectively.

[c] The amplitudes for Strong incongruity were more negative at the right than at the left electrodes ($p<.001$ in each language condition): for Italian right = −3.67 μV and left = −2.60 μV; for jabberwocky right = −2.95 μV and left = −1.72 μV; for French right = −2.26 μV and left = −1.66 μV.

[d] The mean amplitudes at frontal–temporal electrodes were more negative for strong incongruity than for other conditions: Italian = −2.99 μV, −1.19 μV, −1.33 μV; Jabberwocky −2.34 μV, −1.46 μV, −1.08 μV; French −2.10 μV, −.84 μV, −.64 μV, for strong incongruity, weak incongruity and congruous endings, respectively.

[e] The mean amplitudes at parietal electrodes were more positive for strong incongruity than for other conditions: Italian = 7.39 μV, 3.89 μV, 2.47 μV; Jabberwocky = 6.35 μV, 1.95 μV, 1.47 μV; French = 6.12 μV, .72 μV, .76 μV, for strong incongruity, weak incongruity and congruous endings, respectively.

[f] The mean amplitudes at parietal electrodes were more positive for weak incongruity than for other conditions: Italian = 4.53 μV, 7.30 μV, 3.99 μV; Jabberwocky = 3.99 μV, 5.56 μV, 3.79 μV; French = 4.54 μV, 5.09 μV, 3.90 μV, for strong incongruity, weak incongruity and congruous endings, respectively.

* $p<.05$.

2.3. Frequency discrimination threshold (FDT)

Single frequency discrimination thresholds (FDTs, i.e., Δfs) were expressed as a percentage of the fixed tone fundamental

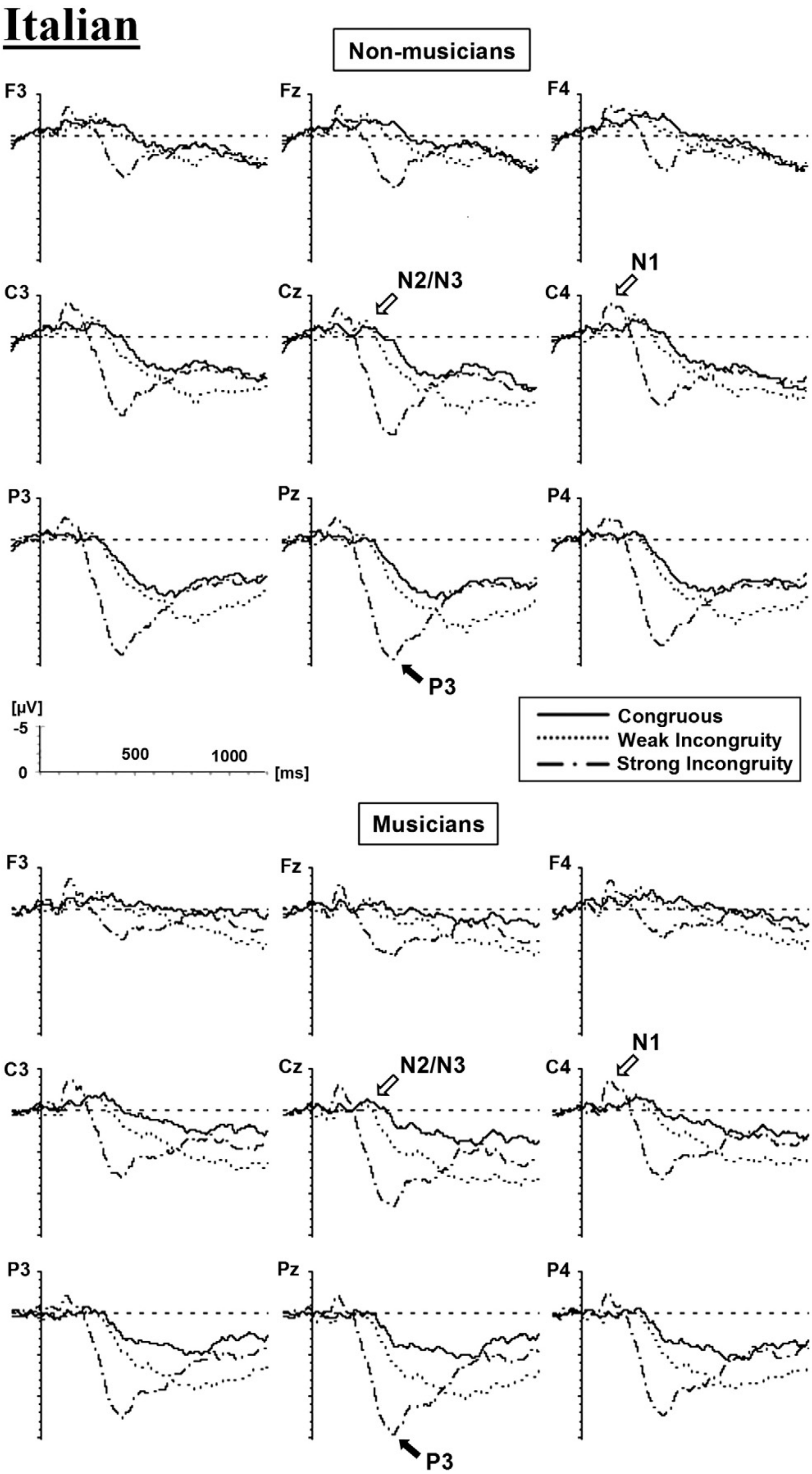

Fig. 3 – Grand-averaged event-related potentials elicited by congruous, weakly incongruous and strongly incongruous endings in Italian sentences obtained from 11 non-musicians (upper panel) and 11 musicians (lower panel). Selected traces from 9 electrodes are presented.

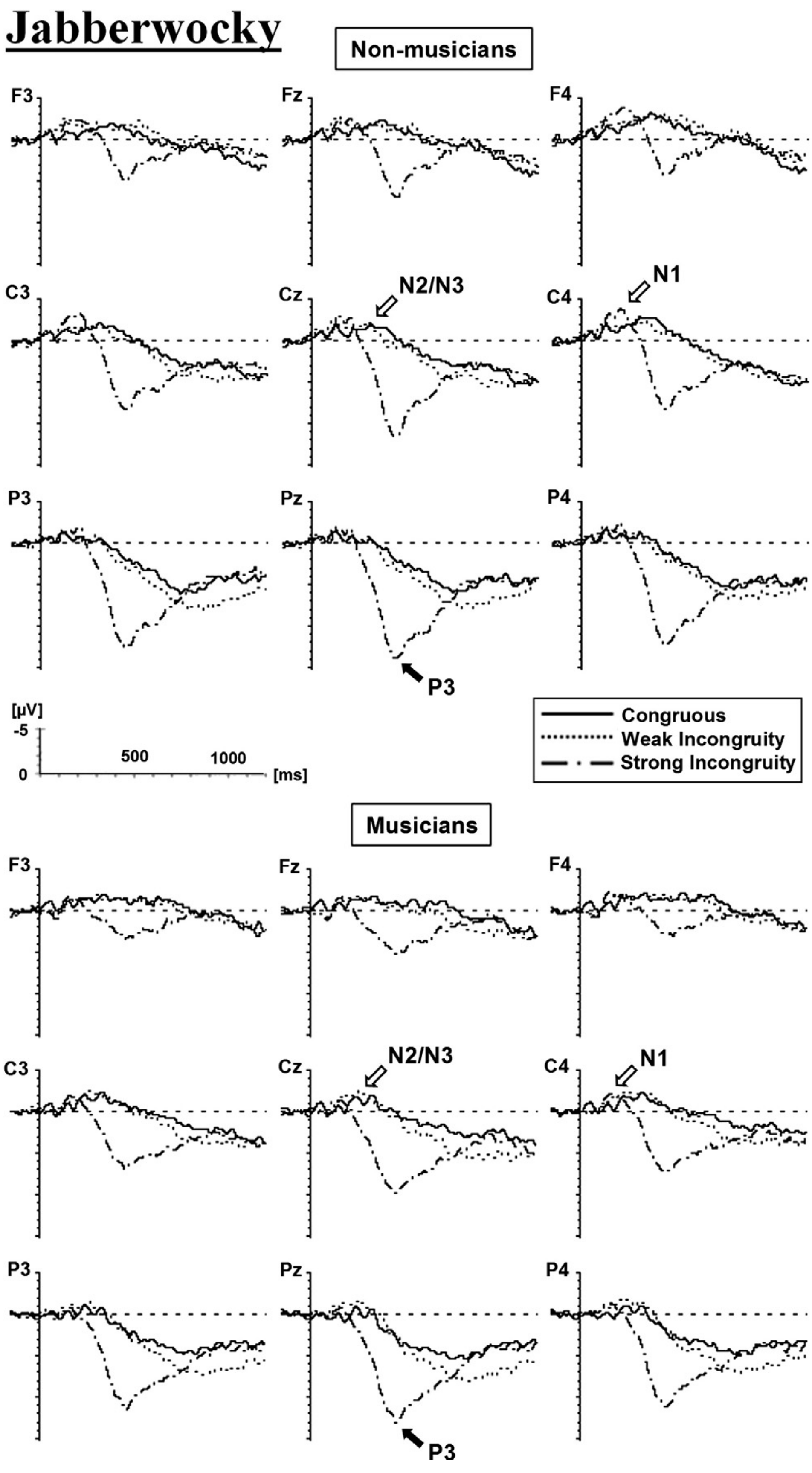

Fig. 4 – Grand-averaged event-related potentials elicited by congruous, weakly incongruous and strongly incongruous endings in jabberwocky sentences obtained from 11 non-musicians (upper panel) and 11 musicians (lower panel). Selected traces from 9 electrodes are presented.

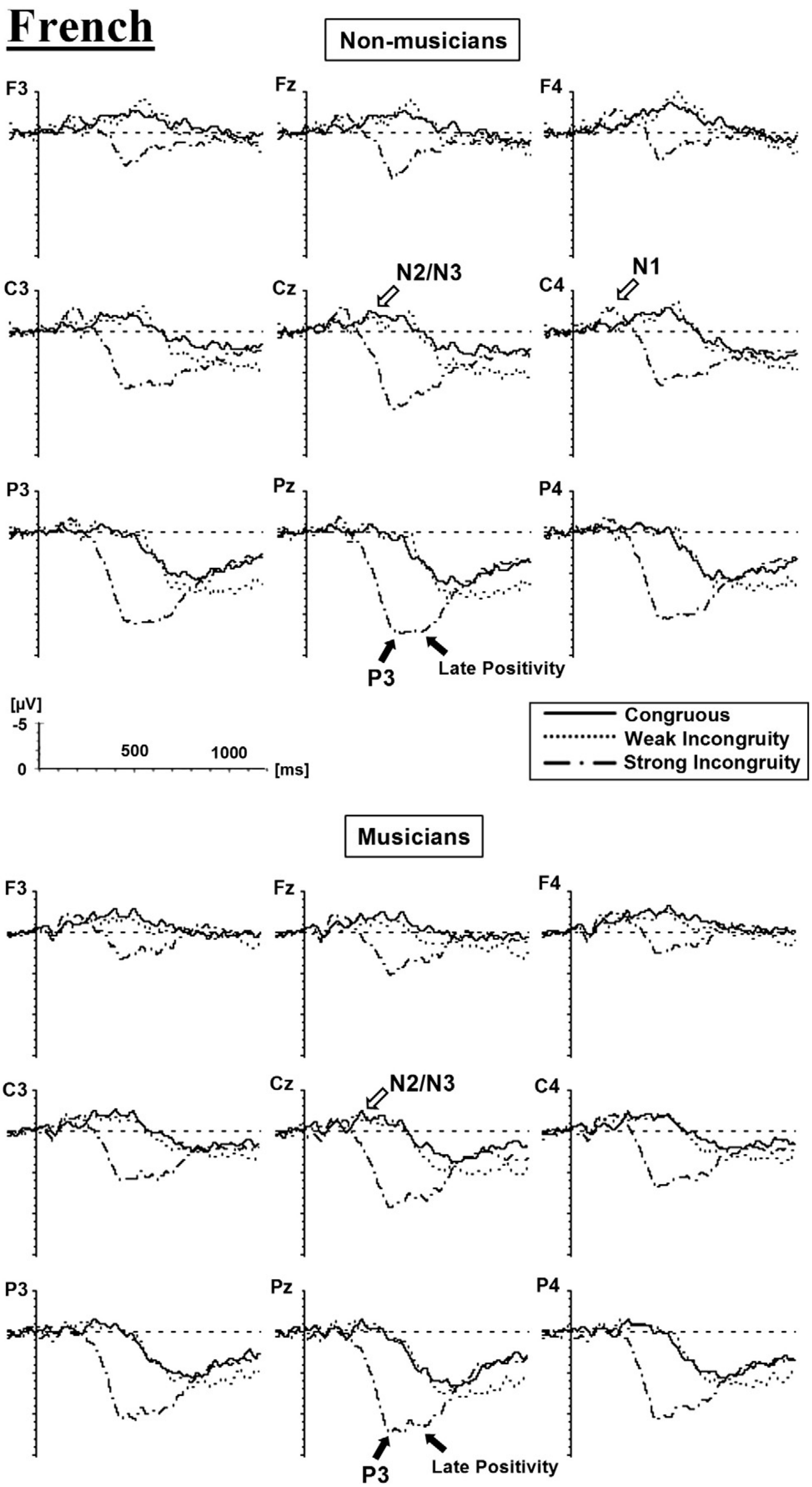

Fig. 5 – Grand-averaged event-related potentials elicited by congruous, weakly incongruous and strongly incongruous endings in French sentences obtained from 11 non-musicians (upper panel) and 11 musicians (lower panel). Selected traces from 9 electrodes are presented.

Table 2 – The onset latency of the positive component at Pz for Weak incongruity determined by comparing with *t*-test the mean ERP amplitudes between Weak incongruity and Congruous condition.

Congruity	Weak incongruity					
Language	Italian		Jabberwocky		French	
Group	Musicians	Non-mus.	Musicians	Non-mus.	Musicians	Non-mus.
350–400 ms	–	–	–	–	–	–
400–450 ms	**	–	–	–	–	–
450–500 ms	**	–	–	–	–	–
500–550 ms	**	–	–	–	–	–
550–600 ms	**	–	–	**	–	–
600–650 ms	**	–	–	–	–	–
650–700 ms	**	–	–	–	–	–
700–750 ms	***	*	–	–	–	–
750–800 ms	***	**	–	–	–	–
800–850 ms	**	**	–	–	–	–
850–900 ms	**	*	*	*	–	–

–: ns.
* $p<.05$.
** $p<.01$.
*** $p<.001$.

frequency for each participant and block. The final FDT of each participant was computed as the geometric mean of the best (i.e., lowest) nine thresholds estimated.[3] For the statistical analysis, log-transformed values of the FDT were used. In line with Micheyl et al. (2006), the mean FDT of musicians (0.24±0.10 [mean lnFDT=−1.52±0.42]) was significantly lower than that of non-musicians (2.05±1.90 [mean lnFDT=.41±0.79]; $t(26)=8.10$, SE=.17, $p<.001$, see Fig. 6). Moreover, non-musicians showed larger within-group variability (range=[0.38: 7.12]) compared to musicians (range=[0.09: 0.43]).

Using FDT threshold as a covariate, further analyses of variance were carried out on error rates in the weak incongruity condition in which sensitivity to subtle perceptual variations were most relevant. If processing differences reflected by the FDT are responsible for the group difference, then when FDT scores are included as covariate no residual effect should remain. In both the error rate and RT analyses a similar pattern was apparent. The Language by Congruity interaction remained significant [for error rate, $F(4,100)=5.96$, MSE=.05, $p<.01$; for RTs, $F(4,100)=3.23$, MSE=29,572, $p<.05$]. By contrast, the group effect disappeared ($F<1$). Thus, a large part of the variance due to the group difference was presumably due to the same factors underlying the listener's sensory acuity in discerning small differences in f0 between harmonic tones, as measured by the FDT task, which is mainly determined by musical training. Had only individual acuity been responsible for the group difference, independently of musical training, we would not have observed such a marked difference between the individuals of the two groups. This does not imply, however, that other factors such as the strategies used by the participants do not play an equally important role.

Regression analyses were carried out with FDT score as predictor and error rates and RTs as dependent variables, in order to see if performance in pitch discrimination for harmonic tones might predict performance on the language task, and if different patterns would emerge for musicians and non-musicians. The analyses were carried out separately for each language and for each group, because of the large between-groups differences in the variance of FDT scores. Error rates to weak incongruities were predicted by FDT scores only in non-musicians, for French sentences, [$r=.64$; $t(13)=2.91$, SE=.07, $p<.05$]: the lower the FDT, the lower the error rate for the weak incongruity in French. FDT did not predict RTs in any congruity or language condition. This result suggests that the two tasks were correlated only when pitch changes in the main task were analyzed on the basis of acoustic information, not on the basis of prosodic or semantic information. The correlation was only present in non-musicians, perhaps due to the small variance in FDTs in musicians.

3. Discussion

In the present paper we investigated the performance of musicians and non-musicians in processing pitch contour in three language conditions. We also analyzed the signal detection indices of sensitivity (d′) and of response criterion (c) calculated from behavioral data in order to evaluate the locus of the musicians' advantage. Further, we analyzed electrophysiological data collected concomitantly with the prosodic pitch task to study the neural substrates underlying perceptual and cognitive processing of pitch changes in the two groups. Finally, we tested the two groups with a psychoacoustic task measuring their threshold to differences in f0 between two non-linguistic tones.

[3] The maximum likelihood procedure is an extremely fast procedure. However, it highly overestimates the threshold if the participant makes an attentional lapse at the very first trial of one block (Grassi and Soranzo, 2009; Green, 1995). Many of our participants were new to experiments in psychoacoustics and occasionally made such attentional lapses. The exclusion of the worst threshold of each participant (in several cases due to such attentional lapses) prevented the estimation of an unrealistic, final threshold.

Table 3 – The onset latency of the positive component at Pz for strong incongruity determined by comparing with *t*-test the mean ERP amplitudes between Strong incongruity and Congruous condition.

Congruity	Strong incongruity					
Language	Italian		Jabberwocky		French	
Group	Musicians	Non-mus.	Musicians	Non-mus.	Musicians	Non-mus.
150–200 ms	–	–	–	–	–	–
200–250 ms	–	–	*	–	–	–
250–300 ms	**	**	***	–	–	–
300–350 ms	**	***	***	**	*	**
350–400 ms	***	***	***	***	**	***

–: ns.
* $p<.05$.
** $p<.01$.
*** $p<.001$.

In line with previous results by Schön et al. (2004) and by Marques et al. (2007) with native and unfamiliar foreign language, respectively, we found that musicians were more accurate than non-musicians when the pitch change was weak and consequently difficult to detect. Analyses of indexes based on the signal detection theory (SDT) also demonstrated higher sensitivity (larger d′) and smaller response bias (smaller c) for musicians than for non-musicians. This advantage persisted when sentences were meaningless (jabberwocky) and when the prosody was unfamiliar (French). However, in contrast to previous results (Marques et al., 2007; Schön et al., 2004), the advantage of musicians was not evident in RTs. This result suggests that when the pitch changes were difficult to detect musicians and non-musicians used different types of strategies. Musicians, being characterized by higher perceptual sensitivity and being trained to evaluate subtle pitch differences, may have slowed down their response to perform a more accurate analysis of the stimulus, before making a decision.[4] Non-musicians, by contrast, responded (relatively) fast, but less accurately to weak pitch changes (i.e., with a larger bias towards "*congruous*" response), not being able to differentiate them from the congruous endings. This difference may reflect that musicians, but not non-musicians, employed a strategy emphasizing accuracy over speed.

ERPs in all language conditions showed a fronto-temporal N1 component to strong incongruities and a late positive component, more pronounced for the strong than for the weak incongruity and congruous endings. Statistical analyses, however, did not reveal effects of musical expertise on the mean amplitude of both components.

Schön et al. (2004) reported different scalp distribution of the N1 component to strong incongruities, which was slightly left-lateralized in non-musicians and more bilaterally distributed in musicians, possibly supporting more efficient auditory processing in musicians. However, this component was not evident in Marques et al. (2007) probably because participants were unsure of the onset of the final words. In the present data the N1 was relatively lateralized to right in both musicians and non-musicians (see Fig. 7) which is compatible with the right-lateralization of processing for sentence pitch contour reported in previous fMRI studies (e.g., Meyer et al., 2002, 2004). This component is known to be susceptible to selective attention (Hillyard et al., 1973), and larger/shorter N1 amplitude/latency to Italian and jabberwocky than to French sentences found by Colombo et al. (2011) can be attributable to higher selective attention to the sentence final word induced by larger expectation generated by familiar prosodic pattern conveyed by Italian and jabberwocky. From this point of view, our results showing no difference in the N1 between musicians and non-musicians are in line with those of Marie et al. (2011, 2012) who concluded that musicians' superiority in pitch processing is not due to more focused selective attention.

Clear effects of musical expertise were revealed by the onset latency of the late positivity to the weak incongruity in the native language, Italian, that was 300 ms shorter for musicians than for non-musicians (see Table 2). Thus, in line with our hypothesis, differences are found when pitch changes are most difficult to detect. This finding is also consistent with

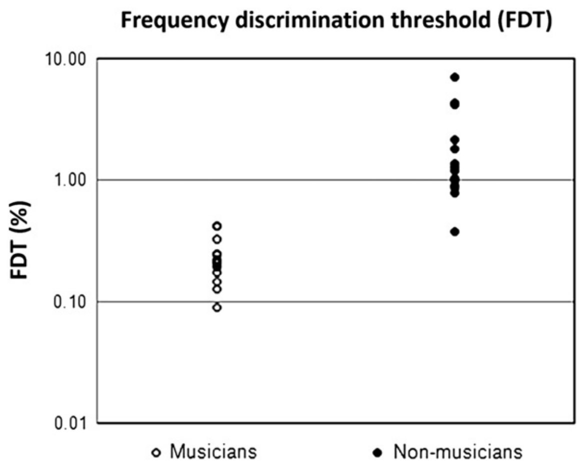

Fig. 6 – Frequency discrimination thresholds of 14 musicians and 14 non-musicians. Each data point represents a single participant.

[4] From the observation of participants' attitude during the experiment (pre-tests included), it was clear that musicians typically wanted to show their improved auditory abilities to process sounds.

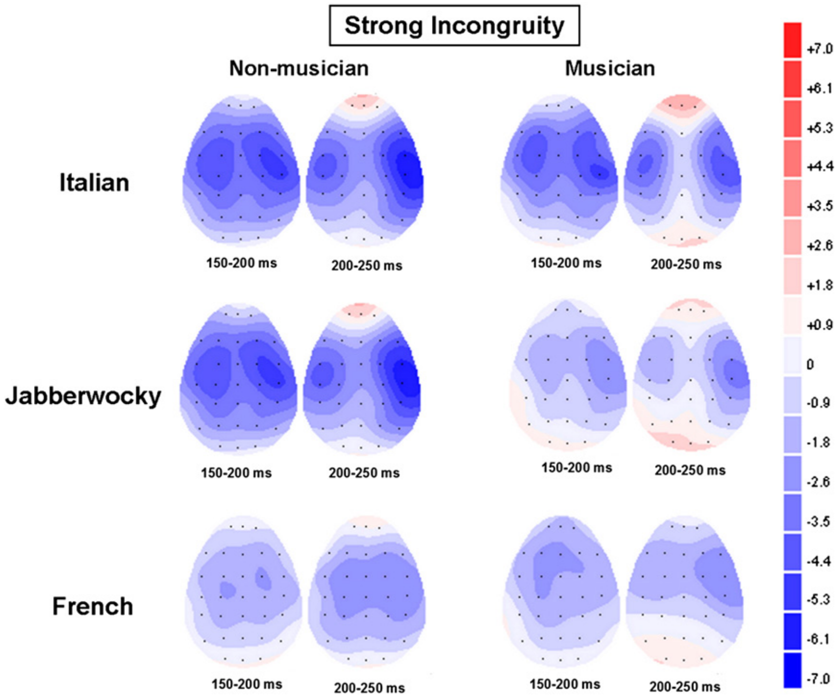

Fig. 7 – Topographic maps of difference in scalp electrical activity between congruous and strongly incongruous conditions obtained from 11 non-musicians and 11 musicians and computed at 50-ms intervals from 150 to 250 ms after the onset of the sentence-last word in Italian, jabberwocky and French.

behavioral data showing that the advantage in accuracy for musicians was particularly strong for weak incongruities in Italian. In this condition, lexical (phonological) and semantic knowledge were likely to facilitate segmentation of the speech stream, and this in turn may have helped detection of the last word of the sentence.

Moreover, the onset latency of the positivity to strong incongruity in jabberwocky sentences was also 100 ms shorter for musicians than for non-musicians and no between-group differences were found for French, possibly because the weak incongruity was too difficult to detect in these contexts. It is interesting to emphasize that in both former studies (Marques et al., 2007; Schön et al., 2004) the evaluation of the weak incongruity was easier, as is apparent from a comparison of accuracy data. For example, error rates for the weak incongruity in the *familiar* language were comparable in the two studies: 13% and 30% in French (Schön et al., 2004) and 14% and 39% in Italian in the present study for musicians and non-musicians, respectively. By contrast, for the weak incongruity in the unfamiliar language, error rates for Portuguese sentences in Marques et al. (2007) were around 15% in musicians and 40% in non-musicians, while in the present study the values for French sentences were 38% and 54%. Thus, the smaller pitch increase (15%) used in the present study possibly produced a larger decrease in accuracy associated with smaller group differences in the electrophysiological data. The failure to completely replicate the former findings, rather than a disadvantage of the present study, suggests important implications for future research. It may suggest that there exists an optimal range of difficulties in which the advantage of musicians becomes apparent. Moreover, different from the former studies, in the present study participants were presented with all the different language conditions, and this may have set up a contextual effect, whereby small pitch differences were less detectable because adaptation to the different stimuli was more complex.

Another potentially relevant issue concerns the effect of participants' native language. Recent studies that addressed the interaction of musical expertise and linguistic experience in auditory processing suggest that native speakers of a tone-language (Bidelman et al., 2011) or a quantity-language (Marie et al., 2012) may have enhanced sensitivity to acoustic changes in specific dimensions that are relevant in their native language (i.e., rapid pitch change, and syllable duration, respectively) at the subcortical and cortical levels. Based on their results, it cannot be excluded that Italian non-musicians in our experiment were differently sensitive to pitch changes in the final segment of the sentence compared to French native speakers, considering that pitch is one of primary parameters defining lexical stress in Italian. However, it is not possible to examine the complex interaction of pitch perception with different linguistic experiences from our data, as a more controlled cross-linguistic design would be required. Nonetheless, this issue suggests avenues of inquiry for further research.

It is also important to note that the sensory and cognitive processes are not independent of each other, because development of later processes should depend on the output of earlier processing. In other words, cognitive strategies employed in pitch incongruity judgment can differ according to the percept resulting from the sensory processing. In this sense, differences in late ERP components should not be interpreted exclusively as results of conscious cognitive mechanisms.

Results from the psychoacoustic test (FDT) clearly showed a significantly lower discrimination threshold and smaller within-group variability in musicians compared to non-musicians, confirming the results by Micheyl et al. (2006). Musicians were indeed much more skilled than non-musicians in detecting pitch differences in tones. Only one non-musician showed a discrimination threshold in the same range as musicians, while in Micheyl et al. (2006) almost a third of participants showed thresholds in the same range as musicians. Evidence from the FDT task further showed that higher accuracy of musicians was probably based on sensory abilities to discriminate subtle pitch differences. However, the language effect (Italian–jabberwocky–French differences) was not accounted for by this ability, as it persisted when FDT scores were used as a covariate. Moreover, FDT only predicted accuracy in the foreign language condition, where prosodic contours were not supported by familiar patterns stored in the long-term memory and could not help detection of incongruities. Thus, the overall pattern suggests that detection of weak pitch changes was carried out in familiar and unfamiliar languages based on different types of information. Presumably in the French condition participants relied more on acoustic cues for their congruity evaluation, while in the familiar languages they relied on prosodic and/or lexical information as well. Indeed, this is consistent with the pattern of accuracy in the three language conditions, where the most accurate condition was one in which both prosodic and lexical information was available.

Overall, for both musicians and non-musicians detection of pitch changes was easier for Italian than for jabberwocky sentences and most difficult for French, as reflected by both behavioral and electrophysiological data. As noted, participants were better able to anticipate the target word on which a pitch change might appear in Italian, based on complete analysis of the sentence level, rather than in jabberwocky, based on prosody alone, or in French, where acoustic cues were predominantly available. In general, the results do not support the idea that the task was carried out by isolating only the pitch dimension from all other levels of linguistic processing in order to analyze task-relevant pitch differences. Rather, all potential information was used in order to perform the task, including semantic and prosodic information that can modulate the sentence pitch perception in a top–down manner. This is also suggested by the fact that the only reliable correlation of the main task with the FDT task was found for French (but only in non-musicians), where both semantic information and familiar prosodic templates were unavailable. Had evaluation of sentences' pitch contour been generally based purely on frequency discrimination, the correlations with the prosodic task would have been larger, at least in the weak incongruity.

In sum, results of the present study showed that musicians are more sensitive than non-musicians to subtle pitch changes both in non-linguistic tones, as measured by FDTs, and in spoken sentences, as indicated by d′. However, and perhaps most interestingly, results also showed that enhanced perceptual sensitivity in musicians is not the whole story. Analyses of behavioral data suggest that musicians and non-musicians use different strategies to perform the task, with the former emphasizing accuracy over speed. Moreover, non-musicians clearly showed a bias toward negative responses to subtle pitch changes thereby suggesting that they did not distinguish them from congruous endings. Finally, analysis of the ERP data clearly showed that musical expertise exerts a strong influence on the onset latency of the P3 component, which is taken to reflect an influence of musical expertise on cognitive abilities such as decision making and the level of confidence in the decision being made.

4. Experimental procedures

4.1. Methods

4.1.1. Participants

Twenty-eight Italian native speakers, 14 musicians and 14 non-musicians, participated in the experiment. Due to major EEG artifacts six participants had to be removed from the analyses, resulting in 11 musicians and 11 non-musicians being available for ERP analyses, while the data of the whole sample were available for behavioral analyses. The mean age of the 14 musicians was 25.3 years (SD = 3.99) and the mean age of the 14 non-musicians was 25.9 years (SD = 5.53). The mean age of the 11 musicians was 24.6 years (SD = 3.67) and the mean age of the 11 non-musicians was 26.4 years (SD = 6.07). Musicians had at least 8 years of training with musical instruments (mean: 13.9 years [SD = 4.46] for 14 musicians and 13.4 years [SD = 4.65] for 11 musicians), while non-musicians had none (except the compulsory music courses at school). Musical and biological background of musicians is shown in Table 4. All participants were right-handed and had normal hearing according to self-report. None of them had learned or understood French. They were paid for participation in the experiment that lasted 2.5 h.

4.1.2. Pitch congruity detection task

4.1.2.1. Materials. Ninety-nine sentences were created for each language condition. The French sentences were selected from the 120 declarative sentences ending with a disyllabic word used by Schön et al. (2004). The Italian sentences were

Table 4 – Background characteristics of 14 musicians who participated in the experiment.

Participant	Sex	Age	Instrument	Years of training
1	M	24	Double bass	8
2	F	21	Violin	15
3	M	27	Clarinet	15
4	M	33	Double bass	20
5	M	34	Oboe	20
6	M	26	Double bass	13
7	M	23	Viola	12
8	M	25	Violin	20
9	M	26	Violin	19
10	M	23	Piano	13
11	M	25	Organ	12
12	F	21	Piano	9
13	F	25	Oboe	12
14	M	21	Double bass	7

similar in content to the French sentences and also ended with a disyllabic word, starting with a stop consonant and stressed on the penultimate syllable (as the majority of Italian words). Jabberwocky sentences were created by substituting the content words of the Italian sentences with pseudowords with the same number of syllables. Examples of sentences in the three language conditions are presented below:

> Italian: *Marco non deve andare in bicicletta sulla corsia riservata ai taxi.*
> Jabberwocky: *Forvo non dusna fadare in chirinetta sulla gornea tirubata ai tessi.*
> French: *Le père de Sophie est accueilli à son arrivée par un délicieux gratin.*

All sentences were recorded by an Italian–French bilingual female speaker. Italian and French sentences were read with their own natural intonation. Each jabberwocky sentence was recorded immediately after the corresponding Italian sentence, in order to help the speaker to maintain a similar prosodic contour in both conditions. Mean sentence duration was 3909 ms (SD=509) for Italian, 4393 ms (SD=612) for jabberwocky and 4104 ms (SD=537) for French. Mean duration of the final word was 507 ms (SD=60) for Italian, 535 ms (SD=61) for jabberwocky, and 511 ms (SD=127) for French. There was no significant difference between languages in the final-word length.

The pitch of the final words of each sentence was manipulated so that it was prosodically congruous, weakly incongruous, or strongly incongruous with the preceding context (see Fig. 8 for an example spectrogram of the sentence). To create the incongruous conditions, the f0 of the (whole) last word was increased by 15% for the weak incongruity and by 100% for the strong incongruity using the software Praat (Boersma and Weenink, 1996). These values were based on the results of pretests, so that accuracy levels in Italian would be comparable to those obtained by Schön et al. (2004) in French.

The 891 stimuli (99 sentences × 3 languages × 3 congruity conditions) were divided into three lists, each containing 33 sentences in each experimental condition. Each participant received one list with 297 stimuli (33 sentences × 3 languages × 3 congruity conditions) so that each sentence was only presented in one congruity condition for each participant but sentences were presented in each congruity condition across participants.

4.1.2.2. Procedure. The experiment took place in an electrically shielded quiet room, and was programmed and presented using the software E-prime (Psychology Software Tools, Pittsburgh, PA). Participants were seated in front of a PC monitor and were required to fixate an asterisk at the centre of the screen while attentively listening to the sentences that were presented via a pair of loudspeakers. They were instructed to decide as quickly and accurately as possible whether the last word of each sentence had a congruous intonation by pressing one of two response buttons (one for "congruous" and the other one for weakly or strongly "incongruous"). For each participant, the three language conditions were presented separately in three experimental blocks. Each experimental block was preceded by a practice block of 9 trials with a feedback to familiarize participants with the task and to train them to blink only during the inter-trial interval. Feedback was only given on practice trials that were not included in the experimental materials. In each block of trials, the sentences were presented in a random order. The hand of response and the order of presentation of the three language conditions were counterbalanced across participants.

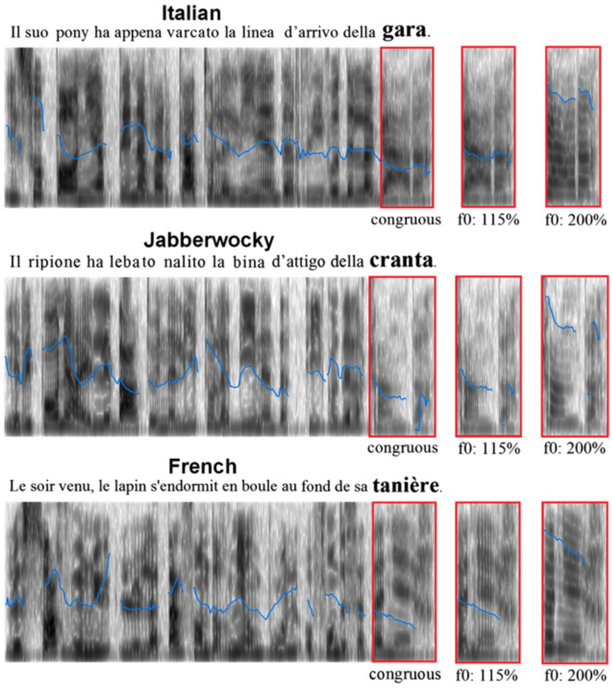

Fig. 8 – Spectrograms of Italian, jabberwocky and French sentences in Congruous, Weakly incongruous and Strongly incongruous conditions. The blue line traces the level of fundamental frequency (f0).

4.1.2.3. ERP recordings. Electroencephalogram (EEG) was recorded from 32 scalp electrodes, mounted on an elastic cap, and located at the following sites according to the International 10/20 system (Jasper, 1958): Fz, Cz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, FCz, CPz, Cp3, Cp4, Fc3, Fc4, TP7, TP8, FPz, FT7, FT8, A2. Data were referenced off-line to the arithmetic mean of the left and right mastoids. Impedances of the electrodes never exceeded 9 kΩ. To detect blinks and vertical eye movements, the horizontal electro-oculogram (EOG) was recorded from electrodes placed 1 cm to the left and right of the external canthi, and the vertical EOG was recorded from electrodes above and beneath the right eye. Trials containing ocular or movement artifacts were excluded from the averaged ERP waveforms. The EEG and EOG were amplified by a Neuroscan amplifier (Compumedics Neuroscan, El Paso, TX) with a bandpass filter of .05–100 Hz and a 50 Hz notch filter, and were digitized at 500 Hz (resolution: .084 μV/LSB).

4.1.2.4. EEG data analysis. ERP data were analyzed by computing mean amplitudes, relative to a 150-ms baseline. Latency windows were selected based upon visual inspection of the waveforms and on preliminary 50 ms latency band analysis between 0 ms and 1200 ms after the onset of the sentence-final word. When the effects were similar in successive 50-ms windows, latency bands were pooled together. ANOVAs were separately computed for midline and lateral electrodes and for each language condition (Italian, jabberwocky and French). They included Expertise (musicians, non-musicians) as between-subject factor and Congruity (congruous, weakly incongruous, and strongly incongruous) as within-subject factor. Midline analyses comprised an Electrode factor (Fz, Cz and Pz), while Lateral analyses comprised Hemisphere (2), ROI (3 regions of interest: fronto-central, temporal, and parietal) and Electrode (3) as within-subject factors. ROIs included: F3, F7, FC3 and F4, F8, FC4; C3, T3, TP7 and C4, T4, TP8; CP3, P3, T5 and CP4, P4, T6. P values were adjusted with the Greenhouse–Geisser epsilon correction for non-sphericity when needed. Tukey tests were used for all post-hoc comparisons. Topographic maps were computed using Neuroscan Edit software (Compumedics Neuroscan, El Paso, TX).

4.1.3. *Psychoacoustic task*

4.1.3.1. Materials. Harmonic complex tones, including the first five harmonics of a 330-Hz (i.e., E4) fundamental frequency, were synthesized at 44,100-Hz and 16 bits resolution. All harmonics were identical in amplitude and phase. All tones presented in the experiment were gated on and off by 10-ms long raised cosine ramps and were presented binaurally via headphones (Sennheiser HD 580) at 65-dB SPL. Frequency discrimination threshold (FDT) were estimated with a classic adaptive procedure (Green, 1990) implemented in a freely downloadable MATLAB toolbox (Grassi and Soranzo, 2009).

4.1.3.2. Procedure. Participants' FDT was estimated in ten blocks of 30 trials each. In each block, the FDT was estimated with the maximum likelihood procedure (Green, 1990) with a 2I-AFC task. In each trial, participants listened to two tones of 200-ms, separated by a 500-ms silent interval. While the fundamental frequency of one tone did not change across trials (330 Hz), that of the other tone varied by a factor Δf. Δf was varied adaptively as a function of the participant's response and according to the maximum likelihood algorithm. Feedback to the response ("*correct*" or "*wrong*" appearing on the screen) was given in each trial. The order of fixed and variable tones was random within each trial. The threshold was taken as the maximum likelihood threshold on the last trial of each block.

Acknowledgments

This study was supported by a grant PRIN 2005 of the Italian Ministry of Instruction, University and Research, and by a grant Progetti di Ateneo 2007 of the University of Padova. The authors wish to thank Marianna Munafò for her assistance in ERP data analyses and Sandro Bettella for technical assistance.

REFERENCES

Banel, M.H., Bacri, N., 1994. On metrical patterns and lexical parsing in French. Speech Commun. 15, 115–126.

Banel, M.H., Bacri, N., 1997. Rôle des indices métriques et des indices phonotactiques lors de la segmentation lexicale en fançais. Annee Psychol. 97, 77–112.

Bidelman, G.M., Gandour, J.T., Krishnan, A., 2011. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J. Cogn. Neurosci. 23, 425–434.

Boersma, P., Weenink, D., 1996. Praat, a system for doing phonetics by computer, version 3.4. Institute of Phonetic Sciences of the University of Amsterdam, Report, 132.

Caniparoli, V., 2001. The role of rhythmic and distributional cues in speech recognition. Ital. J. Linguist. (Rivista di linguistica) 13 (1), 69–84.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., Kraus, N., 2009. Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. Neuron 64 (3), 311–319.

Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47, 236–246.

Colombo, L., Deguchi, C., Boureux, M., Sarlo, M., Besson, M., 2011. Detection of pitch violations depends upon the familiarity of intonational contour of sentences. Cortex 47, 557–568.

Delattre, P., 1947. La liaison en franc¸ais: tendances et classification. Fr. Rev. 21, 148–157.

Delattre, P., 1966. Studies in French and Comparative Phonetics. Mouton, The Hague.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11, 357–374.

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. Audiol. Neurootol. 7, 71–99.

Escera, C., Alho, K., Schröger, E., Winkler, I., 2000. Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol. Neurootol. 5, 151–166.

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.

Grassi, M., Soranzo, A., 2009. MLP: a MATLAB toolbox for rapid and reliable auditory threshold estimations. Behav. Res. Meth. 41 (1), 20–28.

Green, D.M., 1990. Stimulus selection in adaptive psychophysical procedures. J. Acoust. Soc. Am. 87, 2662–2674.

Green, D.M., 1995. Maximum-likelihood procedures and the inattentive observer. J. Acoust. Soc. Am. 97, 3749–3760.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New York, NY.

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.

Ingram, J.G., 1970. Individual differences in signal detection. Acta Psychol. 34, 39–50.

Jasper, H.H., 1958. The ten-twenty electrode system of the International Federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–375.

Kishon-Rabin, L., Amir, O., Vexler, Y., Zaltz, Y., 2001. Pitch discrimination: are professional musicians better than non-musicians? J. Basic Clin. Physiol. Pharmacol. 12 (2 Suppl.), 125–143.

Macmillan, N.A., 1993. Signal detection theory as data analysis method and psychological decision model. In: Keren, G., Lewis, C. (Eds.), A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues. Erlbaum, Hillsdale, NJ, pp. 21–57.

Marie, C., Delogu, F., Lampis, G., Olivetti Belardinelli, M., Besson, M., 2011. Influence of musical expertise segmental and tonal processing in Mandarin Chinese. J. Cogn. Neurosci. 23 (10), 2701–2715.
Marie, C., Kujala, T., Besson, M., 2012. Musical and linguistic expertise influence preattentive and attentive processing of non-speech sounds. Cortex 48 (4), 447–457.
Marques, C., Moreno, S., Castro, S.L., Besson, M., 2007. Musicians detect pitch violation in a foreign language better than non-musicians: behavioral and electrophysiological evidence. J. Cogn. Neurosci. 19 (9), 1453–1463.
McNicol, D., 1972. A Primer of Signal Detection Theory. Allen & Unwin, London.
Meyer, M., Alter, K., Friederici, A.D., Lohmann, G., von Cramon, D.Y., 2002. Functional MRI reveals brain regions mediating slow prosodic manipulations of spoken sentences. Hum. Brain Mapp. 17, 73–88.
Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D., von Cramon, D.Y., 2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang. 89, 277–289.
Micheyl, C., Delhommeau, K., Perrot, J., Oxenham, A.J., 2006. Influence of musical and psychoacoustical training on pitch discrimination. Hear. Res. 219, 36–47.
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. Cereb. Cortex 19, 712–723.
Musacchia, G., Strait, D., Kraus, N., 2008. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear. Res. 241 (1–2), 34–42.
Picton, T.W., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9, 456–479.
Rugg, M.D., Coles, M.G.H., 1995. The ERP and cognitive psychology: conceptual issues. In: Rugg, M.D., Coles, M.G.H. (Eds.), Electrophysiology of Mind: Event Related Brain Potentials and Cognition. Oxford University Press, Oxford, pp. 27–39.
Spiegel, M.F., Watson, C.S., 1984. Performance on frequency-discrimination tasks by musicians and non-musicians. J. Acoust. Soc. Am. 76, 1690–1695.
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.
Wong, P.C.M., Skoe, E., Russo, N., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat. Neurosci. 10, 420–422.