Musicians and the Metric Structure of Words

Céline Marie[1], Cyrille Magne[2], and Mireille Besson[1]

Abstract

■ The present study aimed to examine the influence of musical expertise on the metric and semantic aspects of speech processing. In two attentional conditions (metric and semantic tasks), musicians listened to short sentences ending in trisyllabic words that were semantically and/or metrically congruous or incongruous. Both ERPs and behavioral data were analyzed and the results were compared to previous nonmusicians' data. Regarding the processing of meter, results showed that musical expertise influenced the automatic detection of the syllable temporal structure (P200 effect), the integration of metric structure and its influence on word comprehension (N400 effect), as well as the reanalysis of metric violations (P600 and late positivities effects). By contrast, results showed that musical expertise did not influence the semantic level of processing. These results are discussed in terms of transfer of training effects from music to speech processing. ■

INTRODUCTION

Prosody is an essential aspect of speech that conveys both emotional and linguistic information. Here, we focus on the linguistic function of prosody and, specifically, on the role of prosodic cues such as rhythm and meter in speech segmentation (i.e., the segmentation of a continuous flow of sounds into meaningful linguistic units). Rhythm is a structuring element of speech and can be defined as "the temporal organization of prominences or accents" (Astésano, 2001, p. 9). Meter, defined as the perception of "a regular pulse made up of beats" (The New Grove Dictionary of Music and Musicians, p. 222), is considered as an important aspect of the perception of rhythm (Ladinig, Honing, Haden, & Winkler, 2009). Previous results have shown that, in many languages including French, metric accents located at the end of rhythmic groups of words are characterized by a lengthening of the final syllable and contribute to the rhythmic organization of speech (Frazier, Clifton, & Carlson, 2004; Bailly, 1989; Wenk & Wioland, 1982). For instance, French listeners use final syllabic lengthening to speed up detection of a target syllable located at a rhythmic-group boundary in comparison to the same syllable at another location (Christophe, Peperkamp, Pallier, Block, & Mehler, 2004; Dahan, 1996). Moreover, Salverda, Dahan, and McQueen (2003) showed that segmental lengthening of phonemically identical sequences in Dutch (e.g., "ham" monosyllabic word is longer than "ham" in "hamster") strongly influences their lexical interpretations. More generally, both adults and infants are sensitive to the rhythmic structure of their native language (e.g., Weber, Hahne, Friedrich, & Friederici, 2004; Cutler, McQueen, Norris, & Somejuan, 2001; Di Cristo, 1998; Nazzi, Bertoncini, & Mehler, 1998), and infants are able to discriminate between nonnative languages from different rhythmic classes (Nazzi & Ramus, 2003; Ramus, 2002).

Many results in the literature using both behavioral and brain imaging methods have shown increased abilities for complex auditory processing in musicians than in nonmusicians (e.g., Levitin & Menon, 2003; Tillmann, Janata, & Bharucha, 2003; Maess, Koelsch, Gunter, & Friederici, 2001; Münte, Kohlmetz, Nager, & Altenmuller, 2001; Münte, Nager, Rosenthal, Johannes, & Altenmuller, 2000; Koelsch, Schröger, & Tervaniemi, 1999; Besson & Faïta, 1995). Directly related to our concern, musical expertise increases the automatic and controlled perception of the metric and rhythmic structures of music (Jongsma et al., 2005; Jongsma, Desain, & Honing, 2004; Jongsma, Quian Quiroga, & van Rijn, 2004; Rüsseler, Altenmuller, Nager, Kohlmetz, & Münte, 2001; Takegata, Paavilainen, Näätänen, & Winkler, 2001; Drake, Penel, & Bigand, 2000; Clarke, 1999; Besson & Faïta, 1995; Alain, Woods, & Ogawa, 1994), and recent results showed that the effect of musical expertise is larger when participants detected meter changes than rhythmic changes (Geiser, Ziegler, Jancke, & Meyer, 2009). But will this advantage extend to processing the metric structure of words in speech?

To examine the neurophysiological basis of meter processing, and specifically of final syllabic lengthening in speech comprehension, Magne et al. (2007) conducted an experiment in which they created metric incongruities on sentence-final trisyllabic words (Magne et al., 2007). They used a specific time-stretching algorithm to increase the duration of the vowels of the penultimate syllable without modifying their timbre or frequency (Pallone, Boussard, Daudet, Guillemain, & Kronland-Martinet, 1999).

[1]Institut de Neurosciences Cognitives de la Méditerranée CNRS and Aix-Marseille Universités, Marseille, France, [2]Middle Tennessee State University, Murfreesboro, TN

Moreover, to study the relationship between metric and semantic processing, they also used sentences ending with semantically congruous or incongruous words (2-by-2 design). Finally, they examined the influence of attention by asking participants to perform either a prosodic or a semantic task. By recording the variations in brain electrical activity (i.e., event-related potentials, ERPs) elicited by metric and semantic incongruities, Magne et al. demonstrated the on-line processing of the words' metric structure when participants focused their attention on the metric aspect of speech. Indeed, results in the metric task showed both an enhancement of the N400 and P700 components to metric incongruities. Because the N400 enhancement to metric incongruities was also present in the semantic task (while final words were semantically congruous), this result was taken as evidence for an automatic processing of the metric structure of words that impeded lexical access and word comprehension. By contrast, the increase in amplitude of the P700 component to metric incongruities was only found in the metric task, so that the P700 was interpreted as reflecting controlled attentional metric processing. These findings highlighted the complex interactions between semantic and auditory processing of basic acoustic parameters (duration) when embedded in speech stimuli (metrical structure). A recent study by Schmitdt-Kassow and Kotz (2009) used a similar protocol as Magne et al. to investigate the interaction between syntax and meter, and found a similar pattern of results (N400 and P600) to metrical incongruities.

In both the Schmitdt-Kassow and Kotz (2009) and Magne et al. (2007) experiments, all participants were nonmusician adults. Interestingly, previous results have provided evidence for transfer of training from music to speech regarding pitch processing, in both adults (Marques, Moreno, Castro, & Besson, 2007; Schön, Magne, & Besson, 2004) and children (Magne, Schön, & Besson, 2006). Thus, the specific aim of the present study was to test the hypothesis that musicians have increased abilities to process the metric structure of words. To this aim, we used the same materials and design as in the study by Magne et al. We predicted that musicians would discriminate metric incongruities (i.e., incongruous final syllabic lengthening) better than nonmusicians in the metric task (i.e., when they focus their attention on the metric aspects of words). This enhancement in performance should be reflected by an increase in the amplitude of the N400 and P700 components to metric incongruities. Furthermore, if the interpretation proposed by Magne et al. following which violations of words' metric structure impedes word comprehension is correct, increased sensitivity to metric structure in musicians, as recently shown by Geiser et al. (2009), may be even more detrimental to comprehension than in nonmusicians. As a consequence, in the semantic task (i.e., when focusing attention on the semantics of the sentence), musicians should show higher error rates and larger N400 components to metrically incongruous words (independently of whether they are semantically congruous or incongruous) than nonmusicians. However, if, as argued by some authors (Moreno et al., 2008; Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Scott, 1992), musicians have increased abilities to focus attention on the task at hand, results may reveal less interference between the unattended (meter) and the attended dimension (semantic). In other words, the percentage of errors and N400 amplitude should be smaller for musicians than for nonmusicians in all conditions.

METHODS

Participants

Fourteen musicians (8 women, mean age = 26 years, age range = 21–30 years) gave their informed consent, and were paid to participate in the experiment that lasted for about one hour and a half. All were right-handed, native speakers of French, and without hearing or neurological disorders. They started musical practice around the age of 6 years old and they had 17 years of musical practice, on average, at the time of testing (see Table 1). None of the musicians reported having absolute pitch.

Stimuli

To be able to compare results, stimuli and design were exactly the same as in the Magne et al. (2007) study. A total of 512 sentences ending with trisyllabic words (i.e., "*Les*

Table 1. Musical Background of Musicians

Musicians	*Age of Onset (Year)*	*Duration (Years)*	*Instruments*
1	4	22	piano, guitar bass
2	5	18	piano, guitar bass, accordion
3	8	18	clarinet, flute, voice
4	5	12	saxophone, guitar bass, accordion
5	5	18	violin
6	8	20	violin, clarinet, piano
7	6	20	piano, organ
8	6	21	oboe, percussion
9	8	17	double bass
10	10	12	saxophone
11	9	17	guitar; guitar bass
12	6	12	flute, saxophone, percussion
13	7	22	flute, percussion, voice
14	6	15	piano, flute, voice
Mean	6.6	17.43	

musiciens ont composé une mélodie" ["The musicians composed a melody"]) were used. Sentences were read by a male native French speaker and recorded in an anechoic room. To measure final syllabic lengthening, each sentence-final word was segmented in three syllables. Averaged final syllabic lengthening (i.e., third syllable) was 1.7 times longer than the average length of nonstressed syllables (i.e., first and second syllables). This ratio is within the range (1.7–2.1) previously reported for averaged final syllabic lengthening in prosodic groups in French (Astésano, 2001).

Rhythmic incongruities were created by applying this 1.7 ratio to the penultimate syllable of the final sentence word while leaving final syllabic duration unchanged so that the structure of metrically incongruous words was 1–1.7–1.7. A time-stretching algorithm was used to manipulate the duration of the acoustic signal without altering its timbre or frequency (Ystad et al., 2007; Pallone et al., 1999). Semantic incongruities were created by replacing sentence final congruous words by an unexpected word that did not make sense in the sentence context (i.e., *"Les musiciens ont composé une bigamie"* ["The musicians composed a bigamy"]).

By independently manipulating rhythmic and semantic congruency of the trisyllabic final words, four experimental conditions were created with (1) metrically and semantically congruous (M+S+), (2) metrically congruous and semantically incongruous (M+S−), (3) metrically incongruous and semantically congruous (M−S+), and (4) metrically and semantically incongruous (M−S−) sentence-final words.

Procedure

Participants were presented with 128 short sentences that were semantically and/or metrically congruous or incongruous (32 trials for each condition). Sentences were presented aurally, through loudspeakers, in a pseudorandom order, within four blocks. In two blocks, participants were asked to pay attention to semantic content in order to decide whether final sentence words were semantically congruous or incongruous. In the other two blocks, participants were asked to pay attention to syllabic duration in order to decide whether sentence-final words were well-pronounced or not. Participants were required to press one of two buttons as quickly and as accurately as possible to give their response. The side (right or left hand) of the response was balanced across participants. Furthermore, half of the participants began with the semantic task and the other half with the metric task. Finally, congruity was also balanced across participants so that sentences that were metrically and/or semantically congruous for half of the participants and were incongruous for the other half. Four different experimental lists of 128 sentences were built in order to present each sentence in each experimental condition across subjects, with no repetition within subjects. Each experiment began with a practice session to familiarize participants with the task and to train them to blink during the interstimulus interval.

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 electrooculogram 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 a BioSemi amplifier. The EEG was re-referenced off-line to the algebraic average of the left and right mastoids and filtered with a bandpass of 0.1–30 Hz (12 dB/oct). 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 containing ocular artifacts, movement artifacts, or amplifier saturation were excluded from the averaged ERP waveforms.

Data Analyses

Behavioral data (error rates and RTs) were analyzed using a three-way analysis of variance (ANOVA), including task (metric or semantic), meter (congruous vs. incongruous), and semantics (congruous vs. incongruous) as within-subject factors.

Based upon visual inspection of the traces and on previous results, ERPs data were analyzed by computing mean ERP amplitudes in the 150–250 msec, 250–350 msec, 350–450 msec, 500–700 msec, and 700–900 msec latency bands. ANOVAs were conducted for both midline and lateral electrodes. ANOVAs for midline electrodes included task (metric or semantic), meter (congruous vs. incongruous), semantics (congruous vs. incongruous), and electrodes (Fz, Cz, Pz, Oz) as within-subject factors. ANOVAs for lateral electrodes included task, semantic (or metric) congruity, hemisphere, regions of interest [ROIs] (fronto-central: F3, FC1, FC5/F4, FC2, FC6; temporal: C3, T3, T5/C4, T4, T6; and centro-parietal: CP1, CP5, P3/CP2, CP6, P4), and electrodes.

RESULTS

Musicians

Behavioral Data

Results on the percentage of errors (see Figure 1) revealed that regardless of the direction of attention (i.e., metric or semantic task), musicians made fewer errors for metrically congruous (2.5%) than incongruous words

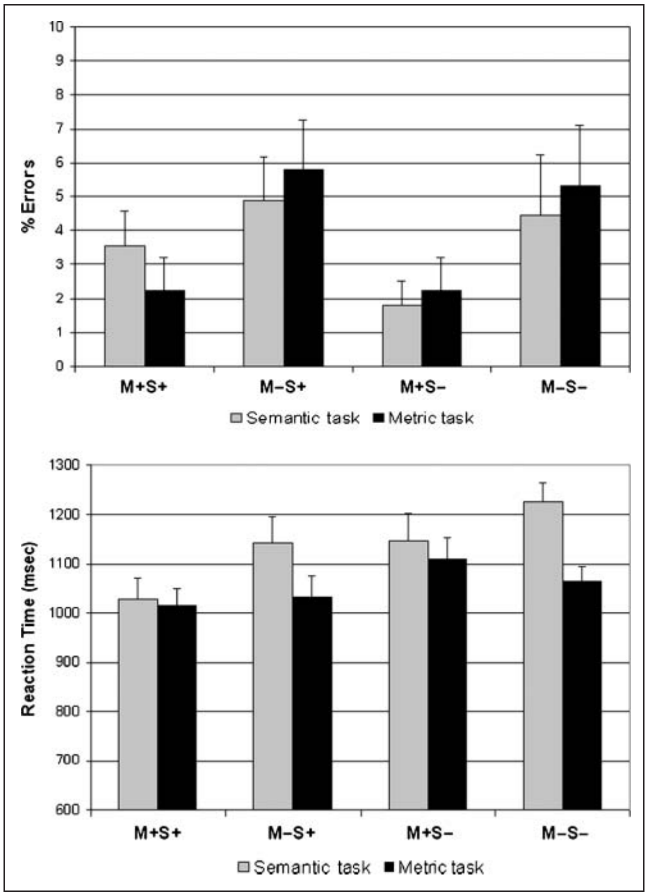


Figure 1. Percentage of errors (top) and reaction times (bottom) for musicians in the semantic (gray) and metric (black) tasks and in the four experimental conditions: M+S+, metrically and semantically congruous; M−S+, metrically incongruous and semantically congruous; M+S−, metrically congruous and semantically incongruous; M−S−, metrically and semantically incongruous sentence-final words.

[5%; main effect of meter: $F(1, 13) = 17.33, p = .001$]. No other effect reached significance.

Results on RTs showed faster responses in the metric task (1055 msec) than in the semantic task [1135 msec; main effect of task: $F(1, 13) = 7.72$, $p = .015$] and for semantically congruous (1075 msec) than incongruous words [1115 msec; main effect of semantics: $F(1, 13) = 21.81$, $p < .001$]. Moreover, musicians' RTs were faster for metrically congruous (1086 msec) than incongruous words (1183 msec) only in the semantic task [$p < .003$; metric task, $R^+ = 1063$ msec and $R^- = 1048$ msec, $p = .89$; Task × Meter interaction: $F(1, 13) = 13.83$, $p = .002$]. No other effect reached significance.

ERP Data

ERPs recorded from musicians and time-locked to the onset of sentence-final words are displayed in Figures 2 and 3. The results of statistical analyses are presented in Table 2. Prior to 150 msec from the final word onset, no significant differences were found either at midline or at lateral electrodes.

Metric congruity effect. Compared to metrically congruous words, metrically incongruous words elicited larger P200 components in the 150–250 msec range (at both midline and lateral electrodes in both tasks: no Task × Meter interaction), larger N400 components in the 350–450 msec range (at lateral electrodes in both tasks: no Task by Meter interaction), larger P600 components in the 500–700 msec range in the metric task only (Task × Meter interaction at both midline and lateral electrodes), and larger late positivities in the 700–900 msec range (at both midline and at lateral electrodes and in both tasks: no Task × Meter interaction). None of the interactions involving the electrodes or ROIs/hemisphere factors were significant in any of the latency bands considered for analyses (see Figure 2).

Semantic congruity effect. Compared to semantically congruous words, semantically incongruous words elicited larger N400 components in the 250–350 msec range at both midline and lateral electrodes and in both tasks. Although the semantic N400 effect seems larger in the semantic than in the metric task, the Task × Semantic interaction was not significant ($p = .87$ at central and $p = .17$ at lateral electrodes). This effect lasted until 900 msec postfinal word onset (see Table 2). None of the interactions involving the electrodes or ROIs/hemisphere factors were significant in any of the latency bands considered for analyses (see Figure 3).

Musicians versus Nonmusicians

To test for the influence of musical expertise on metric and semantic processing during speech comprehension, both behavioral and ERP data from musicians were compared to those from the 14 nonmusicians in the experiment by Magne et al. (2007).

Behavioral Data

Mixed-design ANOVAs, including expertise (musicians vs. nonmusicians), task (metric vs. semantic), meter (congruous vs. incongruous), and semantic (congruous vs. incongruous), were performed on error rates and RTs. Results on error rates revealed that musicians outperformed nonmusicians in both tasks [main effect of expertise: musicians = 3.79% vs. nonmusicians = 17.17%; $F(1, 26) = 184.19$, $p < .0001$]. No other effects reached significance and no significant effects were found on RTs.

ERP Data

To focus on the effects of interest, difference waves were computed separately in both groups by subtracting metrically or semantically congruous endings from incongruous endings (i.e., [(M−) − (M+)] or [(S−) − (S+)]; see Figure 4). Data were analyzed by computing mean amplitudes in the same latency bands as described above. Because the

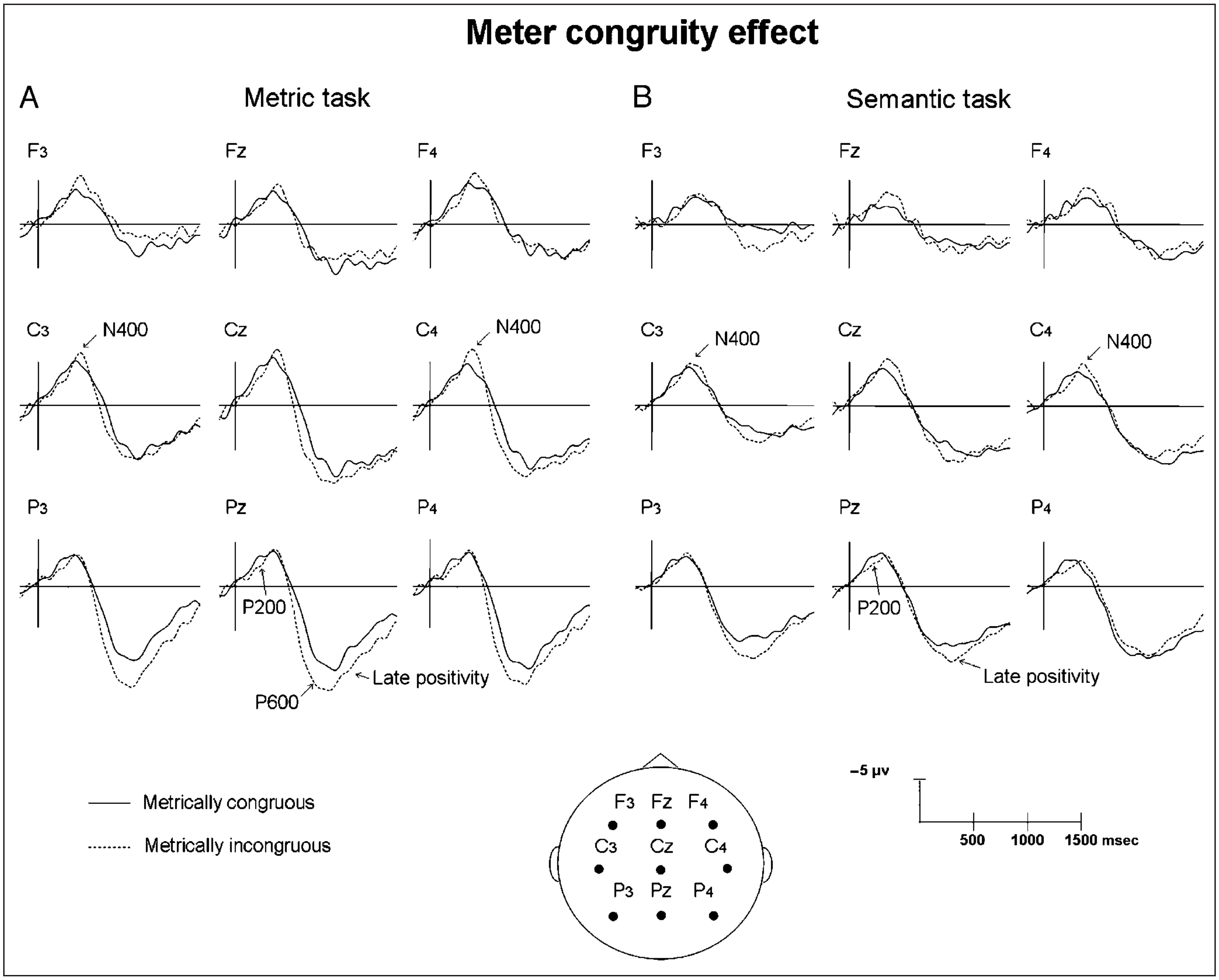


Figure 2. ERPs recorded from musicians and elicited by metrically congruous (solid line) or metrically incongruous (dashed line) final words, in the metric (A) and semantic tasks (B). Selected traces from nine electrodes are presented. In this figure, as in the following ones, the amplitude (μV) is plotted on the ordinate (negative up) and the time (msec) is on the abscissa.

Meter congruity × Semantic congruity interaction was not significant either for musicians or for nonmusicians, the main effects of semantic congruity and meter congruity were analyzed separately. Mixed-design ANOVA included expertise as a between-subject factor together with task (metric vs. semantic) and electrodes (3) for midline analysis, and hemisphere (2), ROI (3), and electrodes (3) for lateral analysis.

Regarding the effect of meter congruity (see Figure 4), the P200 effect in the 150–250 msec range was larger for musicians than for nonmusicians in both tasks at midline and lateral electrodes [$F(1, 26) = 9.08, p < .001$ and $F(1, 26) = 8.64, p = .006$, respectively]. This difference was more pronounced over the right hemisphere [Expertise × Hemisphere interaction: $F(1, 26) = 4.38, p = .046$]. By contrast, the N400 effect in the 350–450 msec range was smaller for musicians than for nonmusicians at lateral electrodes in both tasks [$F(1, 26) = 4.92, p = .03$]. No between-groups differences were found in the P600 latency band (500–700 msec), but the late positivity was slightly larger in musicians than in nonmusicians in both tasks over the right hemisphere in the 700–900 msec latency band [Expertise × Hemisphere : $F(1, 26) = 4.11, p = .05$].

Regarding semantic congruity, no between-groups differences were found in any of the latency bands considered for analysis (see Figure 4).

DISCUSSION

Results for musicians and for the comparison between musicians and nonmusicians are considered, in turn, in the following discussion by considering first the effects of meter and then the semantic effects.

Metric Congruity Effects

The P200 Component

Results revealed that metrically incongruous words elicited larger P2 components than metrically congruous words.

The P200 component is an exogenous component typically considered as reflecting perceptual processing (Hillyard & Picton, 1987). With auditory stimuli, P200 is modulated in amplitude by the acoustic attributes of sounds. For instance, Shahin, Roberts, Pantev, Trainor, and Ross (2005) showed that P2 amplitude is larger when sound spectral complexity (i.e., the number of harmonics in the stimulus) is higher (Shahin et al., 2005). Directly related to our concerns Böcker, Bastiaansen, Vroomen, Brunia, and Gelder (1999) reported larger P2 amplitude to incongruous metric accents when participants passively listened to auditorily presented words (Böcker et al., 1999). In line with these findings, the present P200 effect was elicited in both the metric and the semantic tasks (i.e., independently of the direction of attention) with similar onset latency and scalp distribution (no Task × Metric congruity interaction and no interaction involving electrodes or ROI or hemispheres). Thus, the amplitude enhancement of the P200 component is taken to reflect the automatic processing of the temporal attributes of words and, specifically, of the words' syllabic structure.

Insofar as syllabic lengthening was located on the second syllable, the P200 effect onset latency may seem surprisingly short (significant between 150 and 250 msec post word onset). However, detailed analysis of the duration of the first syllable of each of the 512 trisyllabic words that were used in this experiment showed that while their mean duration was of 150 msec, half of the first syllables were shorter. Thus, for half the words, the second syllable starts earlier than 150 msec (see Figure 5), thereby possibly accounting for the early onset of the P200 effect.

The Effect of Musical Expertise

The P200 effect was larger for musicians than for nonmusicians (see Figure 4). Previous results have shown that musicians are better at preattentively extracting information from musically relevant stimuli (Koelsch et al., 1999) and that extended musical practice influences P200 amplitude in complex auditory pattern discrimination tasks (Shahin et al., 2005; Shahin, Bosnyak, Trainor, & Roberts,

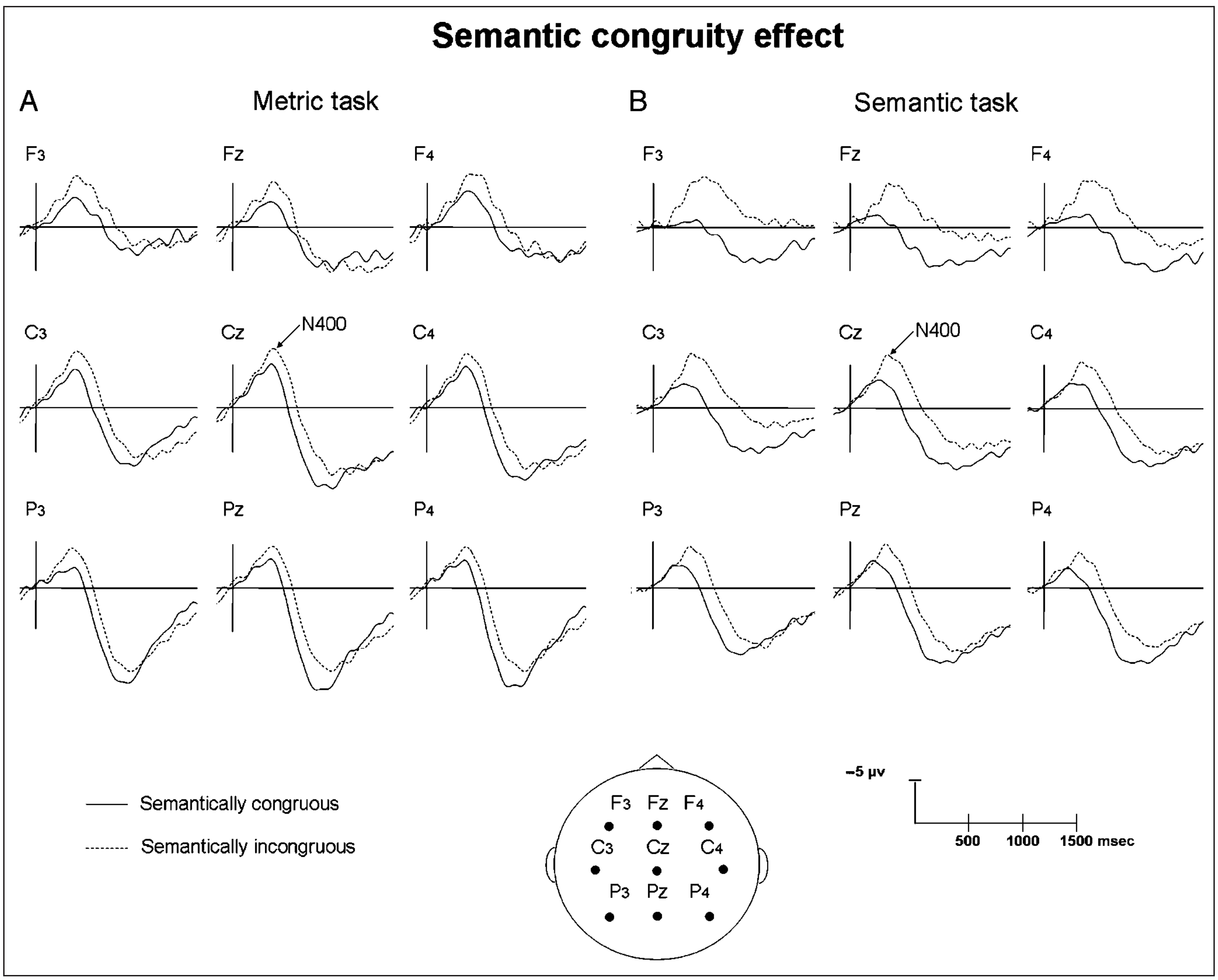


Figure 3. ERPs recorded from musicians and elicited by semantically congruous (solid line) or semantically incongruous (dashed line) final words, in the metric (A) and semantic tasks (B).

Table 2. Summary of Results (ANOVAs) for the Musicians' ERPs Data in Different Latency Bands

		Midline			*Lateral*		
Effects	*Latency Bands (msec)*	*F*	*p*	*d* (μV)	*F*	*p*	*d* (μV)
Metric Congruity Effect							
Meter	0–150	–	–	–	–	–	–
Meter	150–250	$F(1, 13) = 15.76$	.001	0.96	$F(1, 13) = 20.33$	.001	0.84
Meter	250–350	–	–	–	–	–	–
Meter	350–450	–	–	–	$F(1, 13) = 5.36$	.037	−0.72
Task × Meter	500–700	$F(1, 13) = 5.78$	.031	–	$F(1, 13) = 6.62$	.023	–
Meter	700–900	$F(1, 13) = 14.97$	.002	1.26	$F(1, 13) = 5.58$	.034	0.81
Semantic Congruity Effect							
Semantic	0–150	–	–	–	–	–	–
Semantic	150–250	–	–	–	–	–	–
Semantic	250–350	$F(1, 13) = 6.09$	.028	−1.37	$F(1, 13) = 6.86$	.021	−1.37
Semantic	350–450	$F(1, 13) = 17.98$	.001	−2.44	$F(1, 13) = 17.97$	.001	−2.24
Semantic	500–700	$F(1, 13) = 17.50$	.001	−3.3	$F(1, 13) = 17.83$	.001	−3.1
Semantic	700–900	$F(1, 13) = 14.64$	.002	−2.5	$F(1, 13) = 12.94$	.003	−2.5

Factors revealing significant effects were: meter (congruous vs. incongruous) and semantic (congruous vs. incongruous). *F* and *p* values are indicated together with the mean difference (*d*) between conditions (μV).

2003). Moreover, results of recent experiments have shown that short-term auditory training also influences P200 amplitude (Bosnyak, Gander, & Roberts, 2007; Atienza, Cantero, & Dominguez-Marin, 2002; Tremblay & Kraus, 2002; Tremblay, Kraus, McGee, Ponton, & Otis, 2001). For instance, Tremblay et al. (2001) observed an enhancement of the P200 component when nonmusician subjects were trained to discriminate the temporal features of speech signals. Thus, both short- and long-term auditory training improve perception of the temporal properties of speech pattern and induce neuroplastic changes as reflected by modulations of P200 amplitude.

Moreover, the present results also showed that, independently of the direction of attention, musicians were more sensitive to subtle changes in the temporal structure of speech than nonmusicians (the P200 effect was significant in both the metric and semantic tasks for musicians but not for nonmusicians; Magne et al., 2007). These results are taken to reveal transfer of training from music to speech and are in line with those of Reinke, He, Wang, and Alain (2003), showing enhanced P200 with practice in vowel discrimination based on F0. Previously, we also reported that musical expertise improved pitch processing in speech in both adults (Marques et al., 2007; Schön et al., 2004) and children (Magne et al., 2006). In these two studies, however, musical expertise did not influence P200 amplitude, but later components (see below). This may result from the linguistic material in these previous experiments being less well controlled (i.e., more variable final word acoustic characteristics) than here (i.e., trisyllabic final words with stable consonant–vowel–consonant structure).

The N400 Component

In musicians, metrically incongruous words elicited larger negative components than metrically congruous words in the 350–450 msec range. In the study mentioned above, Böcker et al. (1999) also reported an increased negativity (labeled N325) following the P200 component. They considered that the N325 reflected the extraction of metrical stress from the acoustic signal and its transformation into task requirement. However, based on the similar onset latency, morphology, and scalp distribution of the metric and semantic effects found here (see Figures 2 and 3), we tend to consider the increased negativity to metrically incongruous words as an N400 effect. Interestingly, this N400 metric effect was significant in both the metric and semantic tasks, thereby arguing again in favor of the automatic processing of the words' syllabic structures.

However, an alternative interpretation was proposed by Magne et al. (2007), in that violations of words' metric structure may hinder lexical access and word comprehension. In other words, the metric N400 effect would not reflect the automatic processing of metric incongruities, but rather the fact that metric incongruity disrupts access to word meaning and, consequently, elicits an N400

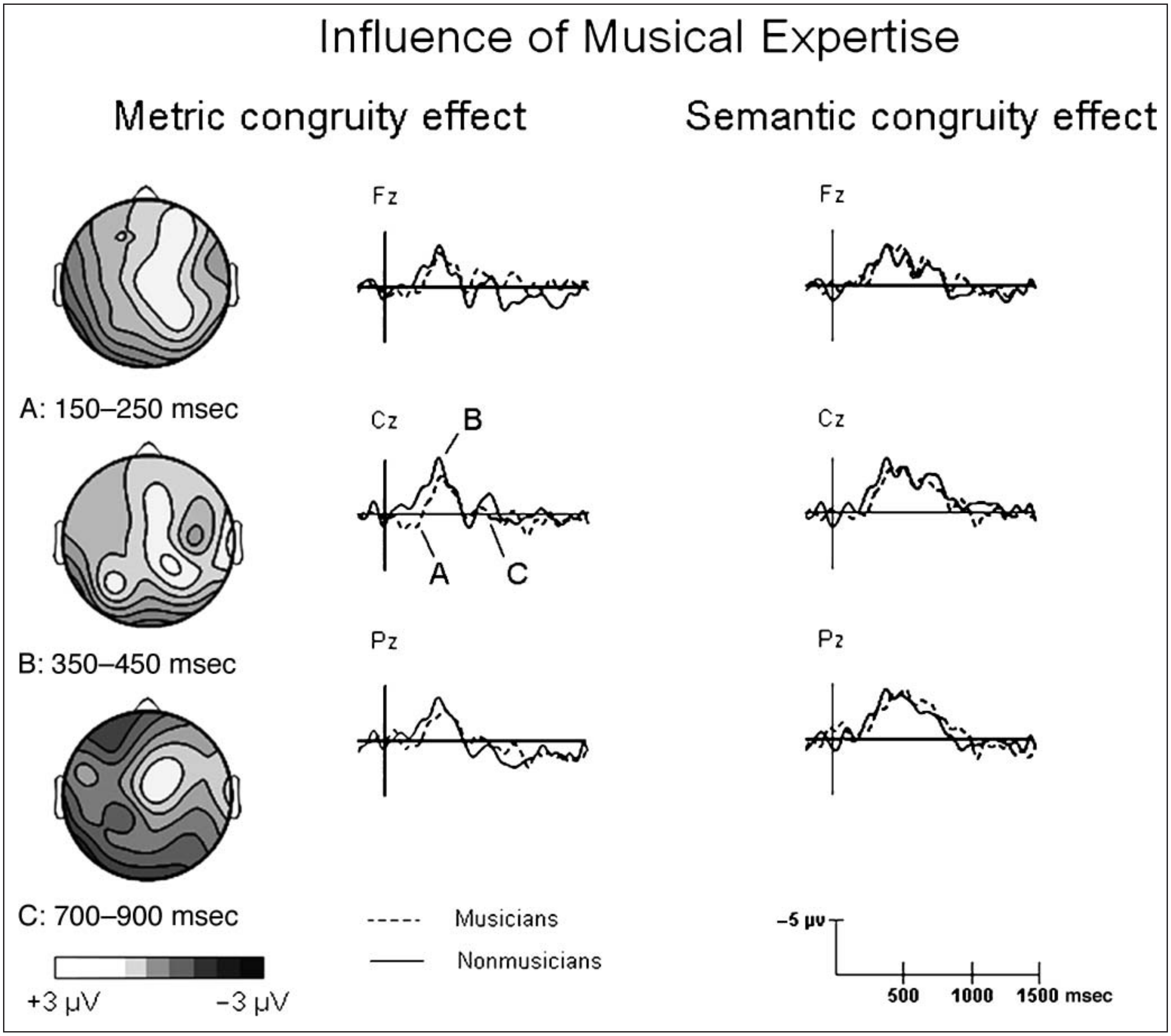
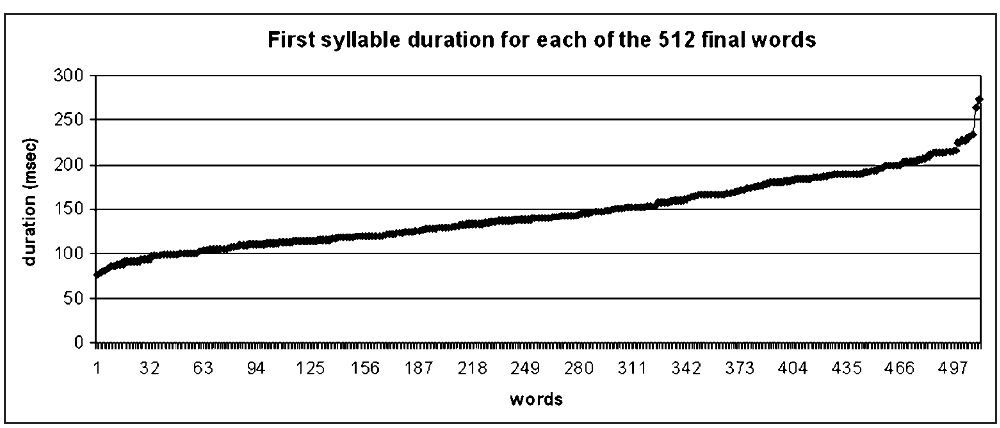


Figure 4. Difference waves (Incongruous minus Congruous conditions) across tasks for the metrical congruity effect (left) and for the semantic congruity effect (right) in musicians (dashed line) and in nonmusicians (solid line). Topographic maps illustrate significant differences between musicians and nonmusicians in three latency bands: (A) P200 (150–250 msec), (B) N400 (350–450 msec), and (C) late positivity (700–900 msec).

Figure 5. First syllable duration for each of the 512 final words.



component. In line with this interpretation, behavioral data in musicians showed that RTs were slower and error rates were higher for metrically incongruous than congruous words in the semantic task (i.e., when participants focused attention on the meaning of the words). Note, however, that these two interpretations do not stand in contradiction because the automatic processing of the words' metric structure may hinder word comprehension.

Recently, Schmidt-Kassow and Kotz (2009) compared metrical and syntactic processing in two attentional conditions (metric and syntactic tasks), in which participants were asked to evaluate the metric homogeneity or the grammatical correctness of sentences. They also found that metrical incongruities elicited a negative component between 250 and 400 msec in both tasks. They considered the two interpretations above, together with the possibility that the negativity is a subcomponent of a left anterior negativity, reflecting the violation of a general (nonlinguistic) rule-based mechanism (Hoen & Dominey, 2000) or a correlate of the ongoing segmentation process. Although these different, but not exclusive, interpretations may account for the N400 metric effects, it is important to note that, independently of the language (German vs. French) and of the specific aspect of meter that was manipulated, results were strikingly similar in both experiments.

The Effect of Musical Expertise

In the Introduction, we reasoned that increased sensitivity to metric structure in musicians than in nonmusicians may be detrimental to comprehension, which could translate into higher error rates and larger N400 components to metrically incongruous words in musicians than in nonmusicians. Alternatively, we also argued that if musicians have increased abilities to focus attention on the task at hand (Moreno et al., 2008; Fujioka et al., 2006; Scott, 1992), the percentage of errors and N400 amplitude should be smaller for musicians than for nonmusicians. Results of the direct comparison between musicians and nonmusicians are in line with this last prediction (see Figure 4). The present results therefore add evidence in favor of increased abilities to selectively focus attention in musicians compared to nonmusicians.

The P600 Component and Late Positivity

Finally, in musicians, metrically incongruous words also elicited larger P600 component than congruous words in the 500–700 msec latency band but only in the metric task. Several interpretations have been proposed for the P600 component, including the proposition that it is syntax-specific and reflects syntactic reanalysis and syntactic integration processes (Friederici, Hahne, & Saddy, 2002; Friederici, Steinhauer, & Frisch, 1999; Hagoort, Brown, & Groothusen, 1993). However, the present results, together with others (e.g., Schmidt-Kassow & Kotz, 2009; Magne et al., 2007), show that the P600 is also sensitive to the metric structure of words or sentences and, consequently, rather reflects a general reanalysis mechanism as proposed by several authors (Eckstein & Friederici, 2005; Magne et al., 2005; Astésano, Besson, & Alter, 2004; Friedrich, Kotz, Friederici, & Alter, 2004; Kaan & Swaab, 2003; Kaan, Harris, Gibson, & Holcomb, 2000). This interpretation is in line with results showing that the P600 is elicited by violations of rule-based expectancies in different domains (e.g., semantic expectancies, Van Herten, Kolk, & Chwilla, 2005; musical expectancies, Magne et al., 2006; Moreno & Besson, 2006; Schön et al., 2004; Besson & Faïta, 1995; metrical expectancies in tone sequences, Abecasis, Brochard, Granot, & Drake, 2005; Brochard, Abecasis, Potter, Ragot, & Drake, 2003).

Interestingly, the present results also showed that although P600 amplitude was larger to incongruous than congruous metric words in the metric task (i.e., when participants were required to make an explicit decision about metric congruity), no P600 differences were found when the same stimuli were presented in the semantic task (i.e., when participants were not required to made an explicit decision on the metric dimension). In this task, differences onset later, in the 700–900 msec range, which is in line with the interpretation that decision processes (Donchin & Coles, 1988; Donchin, 1981) based on stimulus reanalysis onset earlier when attention is directed on the metric dimension than when attention is directed on the semantic dimension (or on the syntactic dimension, as in the study of Schmidt-Kassow & Kotz, 2009).

The Effect of Musical Expertise

In both groups, P600 effects were significant in the metric task but not in the semantic task. These results are in line with the interpretation proposed above following which differences in P600 amplitude reflect general reanalysis processes required to make an explicit categorization and to decide that words were metrically congruous or incongruous.

The most interesting difference between musicians and nonmusicians was that although late positivities (700–900) were only found in the metric task for nonmusicians, they were elicited in both tasks in musicians. We interpret these differences as showing that musicians noticed that the words were metrically congruous or incongruous, independently of the direction of attention. By contrast, nonmusicians only processed metric incongruities when task-relevant. This last result is in line with previous results from Astésano et al. (2004), who tested nonmusicians and showed that prosodic violations (i.e., violations of the intonation pattern of interrogative and declarative sentences) were associated with increased late positivities only when relevant to the task at hand. Taken together, these results again point to the musicians' greater sensitivity than nonmusicians to words' metric structure.

In conclusion, analysis of the metric effects showed that musical expertise influenced the different stages of linguistic metrical processing: the automatic detection of the temporal structure of the syllable (P200 effect); the integration of metric structure and its influence on word comprehension (N400 effect); and the reanalysis of metric violations and controlled decision processes (P600 and late positivities effects). Moreover, by showing increased sensitivity of musicians compared to nonmusicians, they add further evidence in favor of transfer of training effects from music to speech processing.

Semantic Congruity Effects

The N400 Effect

The musicians' semantic congruity effect in the semantic task was characterized by longer RTs and by an increased negativity to semantically incongruous words compared to congruous words (N400 effect) that started around 250 msec postfinal word onset and lasted until 900 msec. This N400 effect was broadly distributed over the scalp (no Congruity × Electrodes or ROI interactions). These results are in line with the literature (Kutas & Hillyard, 1980) and are taken to reflect the greater difficulties encountered in integrating semantically incongruous words in ongoing sentence contexts or in processing words that are unexpected within a given sentence context (DeLong, Urbach, & Kutas, 2005; Kutas & Hillyard, 1984; see also Besson, Magne, & Regnault, 2004 and Kutas & Federmeier, 2000 for reviews).

More revealing is the finding that similar N400 effects were found in both the semantic and metric tasks, thereby arguing for an automatic processing of words' meaning even when participants were instructed to focus attention on syllabic duration. This interpretation is in line with the results of Astésano et al. (2004), showing an N400 to semantic incongruity, independently of whether participants focused their attention on the semantic or prosodic aspects (modality contour, interrogative, or declarative) of the sentence. Although other results also argue for automatic semantic processing (e.g., Heil, Rolke, & Pecchinenda, 2004; Kiefer & Spitzer, 2000; Besson, Fischler, Boaz, & Raney, 1992), other results rather support the attentional control dependency of semantic processing (Chwilla, Brown, & Hagoort, 1995; Brown & Hagoort, 1993). These discrepancies may be linked to differences in task requirements as well as with the modality of stimulus presentation as semantic priming seems to be more automatic in the auditory than in the visual modality (Perrin & Garcia-Larrea, 2003).

The Effect of Musical Expertise

No differences in the ERP semantic congruity effect were found between musicians and nonmusicians, a result in line with those of Bonnel, Faïta, Peretz, and Besson (2001) showing no influence of musical expertise on the detection of semantic incongruities in an explicit detection task. Moreover, Poulain-Charronnat, Bigand, Madurell, and Peereman (2005) also report no influence of musical expertise in an implicit semantic task. Thus, the few results available in the literature show no influence of musical expertise on semantic processing probably because musicians and nonmusicians have similar levels of linguistic expertise. Note, however, that analysis of behavioral data showed that musicians made fewer errors than nonmusicians in all conditions. One interpretation of these seemingly discrepant results (no ERPs effects but difference in behavior) is that, as mentioned above, musical expertise influenced general attentional processes. Learning and practicing music require being able to focus attention and to analyze, in real time, the stream of auditory information to select relevant information and to ignore others. For example, in a symphony orchestra, the violinist needs to share his attention between listening to the sounds that he produces, to the other violins, and to the other instruments of the orchestra. Thus, learning to play music may induce more effective sharing of attentional resources allocated to the task at hand (metric or semantic) with no interference effects between the two types of information (metric and semantic) in the auditory stream. This hypothesis is under current investigation.

Conclusion

In conclusion, the present results showed that musical expertise did not influence the semantic level of processing which is to be expected insofar as both musicians and nonmusicians have similar levels of linguistic abilities. By contrast, musicians processed the metric structure of words better than nonmusicians. Moreover, electrophysiological data revealed that musical expertise increased both the automatic and controlled aspects of auditory processing. These results are taken to show that musicians are more sensitive to sound acoustic parameters not only in music but also in speech. They highlight transfer of training effects from music to speech when considering the metric structure of words, and thereby add evidence in favor of similar processing of meter in music and speech. They also extend previous results showing a positive influence of musical training on linguistic pitch processing (Marquez et al., 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007; Magne et al., 2006; Schön et al., 2004; Reinke et al., 2003).

Acknowledgments

This research was supported by a grant from the ANR-NEURO (ANR-Neuro#024-01) to M. Besson and was conducted at the Institut de Neurosciences Cognitives de la Méditerranée. C. Marie is supported by a fellowship from the French Ministère de l'Education Nationale et de la Recherche.

Reprint requests should be sent to Céline Marie, INCM, CNRS-Marseille, 31-Chemin Joseph Aiguier, Marseille, 13402, France, or via e-mail: Celine.Marie@incm.cnrs-mrs.fr.

REFERENCES

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception, 22,* 549–562.

Alain, C., Woods, D. L., & Ogawa, K. H. (1994). Brain indices of automatic pattern processing. *NeuroReport, 6,* 140–144.

Astésano, C. (2001). *Rythme et accentuation en français. Invariance et variabilité stylistique*. Paris: L'Harmattan.

Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research, 18,* 172–184.

Atienza, M., Cantero, J. L., & Dominguez-Marin, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning & Memory, 9,* 138–150.

Bailly, G. (1989). Integration of rhythmic and syntactic constraints in a model of generation of French prosody. *Speech Communication, 8,* 137–146.

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., Fischler, I., Boaz, T., & Raney, G. (1992). Effects of automatic associative activation on explicit and implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 89–105.

Besson, M., Magne, C., & Regnault, P. (2004). Le traitement du langage. In B. Renault (Ed.), *L'imagerie fonctionnelle électrique (EEG) et magnétique (MEG): Ses applications en sciences cognitives*. Paris: Hermès.

Böcker, K. B., Bastiaansen, M. C., Vroomen, J., Brunia, C. H., & Gelder, B. D. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology, 36,* 706–720.

Bonnel, A. M., Faïta, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics, 63,* 1201–1213.

Bosnyak, D. J., Gander, P. E., & Roberts, L. E. (2007). Does auditory discrimination training modify representations in both primary and secondary auditory cortex? *International Congress Series, 1300,* 25–28.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14,* 362–366.

Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5,* 34–44.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language, 51,* 523–547.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32,* 274–285.

Clarke, E. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 473–500). San Diego, CA: Academic Press.

Cutler, A., McQueen, J. M., Norris, D., & Somejuan, A. (2001). The roll of the silly ball. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 181–194). Cambridge, MA: MIT Press.

Dahan, D. (1996). *The role of rhythmic groups in the segmentation of continuous French speech*. Proceedings of the Fourth International Conference of Speech and Language Processing; October 3–6; Philadelphia, pp. 1185–1188.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117–1121.

Di Cristo, A. (1998). French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195–218). Cambridge, UK: Cambridge University Press.

Donchin, E. (1981). Surprise! Surprise? *Psychophysiology, 18,* 493–513.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11,* 355–372.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18,* 1–23.

Eckstein, K., & Friederici, A. D. (2005). Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research, 25,* 130–143.

Frazier, L., Clifton, C., & Carlson, K. (2004). Don't break, or do: Prosodic boundary preferences. *Lingua, 114,* 3–27.

Friederici, A., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.

Friederici, A., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition, 27,* 438–453.

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cognitive Brain Research, 20,* 300–308.

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.

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

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–483.

Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic semantic activation is no myth: N400 semantic priming effects in the letter search task in the absence of RT effects. *Psychological Science, 15,* 852–885.

Hillyard, S. A., & Picton, T. W. (1987). Electrophysiology of cognition. In F. Plum (Ed.), *Handbook of physiology: Section 1. The nervous system. Volume V. Higher functions of the brain, part 2* (pp. 519–584). Bethesda, MD: American Physiological Society.

Hoen, M., & Dominey, P. (2000). ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *NeuroReport, 11,* 3187–3191.

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. *Biological Psychology, 66,* 129–152.

Jongsma, M. L. A., Quian Quiroga, R., & van Rijn, C. M. (2004). Rhythmic training decreases latency jitter of omission evoked potentials (OEPs). *Neuroscience Letters, 355,* 189–192.

Jongsma, M. L., Eichele, T., Quian Quiroga, R., Jenks, K. M., Desain, P., Honing, H., et al. (2005). Expectancy effects on omission evoked potentials in musicians and non musicians. *Psychophysiology, 42,* 191–201.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15,* 159–201.

Kaan, E., & Swaab, T. Y. (2003). Electrophysiological evidence for serial processing: A comparison between nonpreferred and ungrammatical continuations. *Cognitive Brain Research, 17,* 621–635.

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *NeuroReport, 11,* 2401–2407.

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

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4,* 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.

Ladinig, O., Honing, H., Haden, G., & Winkler, I. (2009). Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception, 26,* 377–386.

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: A possible role for Brodmann area 47 in temporal coherence. *Neuroimage, 20,* 2142–2152.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4,* 540–545.

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: Behavioral and electrophysiological evidence. *Cerebral Cortex, 17,* 2659–2668.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., & Besson, M. (2005). On-line processing of "pop-out" words in spoken French dialogues. *Journal of Cognitive Neuroscience, 17,* 740–756.

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.

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.

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. (2008). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19,* 712–723.

Münte, T. F., Kohlmetz, C., Nager, W., & Altenmuller, E. (2001). Neuroperception. Superior auditory spatial tuning in conductors. *Nature, 409,* 580.

Münte, T. F., Nager, W., Rosenthal, O., Johannes, S., & Altenmuller, E. (2000). Attention to pitch in musicians and nonmusicians: An event-related brain potential study. In T. Nakada (Ed.), *Integrated human brain science* (pp. 389–398). Amsterdam: Elsevier.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination in newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 756–777.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication, 41,* 233–243.

Pallone, G., Boussard, P., Daudet, L., Guillemain, P., & Kronland-Martinet, R. (1999). *A wavelet based method for audio video synchronization in broadcasting applications*. In Proceedings of the DAFX'99, Trondheim, Norway, (December 9–11).

Perrin, F., & Garcia-Larrea, L. (2003). Modulation of the N400 potential during auditory phonological/semantic interaction. *Cognitive Brain Research, 17,* 36–47.

Poulain-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition, 94,* B67–B78.

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition, 2,* 85–115.

Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research, 17,* 781–791.

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

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90,* 51–89.

Schmitdt-Kassow, M., & Kotz, S. (2009). Event-related potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience, 21,* 1693–1708.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music facilitates pitch processing in language. *Psychophysiology, 41,* 341–349.

Scott, L. (1992). Attention and perseverance behaviors of preschool children enrolled in Suzuki violin lessons and other activities. *Journal of Research in Music Education, 40,* 225–235.

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in skilled musicians. *Journal of Neuroscience, 23,* 5545–5552.

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *NeuroReport, 16,* 1781–1785.

Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (2001). Preattentive processing of spectral, temporal, and structural characteristics of acoustic regularities: A mismatch negativity study. *Psychophysiology, 38,* 92–98.

Sadie, S. (Ed.). *The New Grove Dictionary of Music and Musicians* (20 volumes). London: Macmillan.

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16,* 145–161.

Tremblay, K., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research, 45,* 564–572.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1–P2 complex after speech–sound training. *Ear and Hearing, 22,* 79–90.

Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22,* 241–255.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of words stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research, 18,* 149–161.

Wenk, B. J., & Wioland, F. (1982). Is French really syllable timed? *Journal of Phonetics, 10,* 193–216.

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 Brief Communication, 10,* 420–422.

Ystad, S., Magne, M., Farner, S., Pallone, G., Aramaki, M., Besson, M., et al. (2007). Electrophysiological study of algorithmically processed metric/rhythmic variations in language and music. *EURASIP Journal on Audio, Speech, and Music Processing, 2007,* 1–14, Article ID 30194.