

Musician Children Detect Pitch Violations in Both Music and Language Better than Nonmusician Children: Behavioral and Electrophysiological Approaches

Cyrille Magne^{1,2}, Daniele Schön^{1,2}, and Mireille Besson^{1,2}

Abstract

■ The idea that extensive musical training can influence processing in cognitive domains other than music has received considerable attention from the educational system and the media. Here we analyzed behavioral data and recorded event-related brain potentials (ERPs) from 8-year-old children to test the hypothesis that musical training facilitates pitch processing not only in music but also in language. We used a parametric manipulation of pitch so that the final notes or words of musical phrases or sentences were congruous, weakly incongruous, or strongly incongruous. Musician children outperformed nonmusician children in the detection of the weak incongruity in both music and lan-

guage. Moreover, the greatest differences in the ERPs of musician and nonmusician children were also found for the weak incongruity: whereas for musician children, early negative components developed in music and late positive components in language, no such components were found for nonmusician children. Finally, comparison of these results with previous ones from adults suggests that some aspects of pitch processing are in effect earlier in music than in language. Thus, the present results reveal positive transfer effects between cognitive domains and shed light on the time course and neural basis of the development of prosodic and melodic processing. ■

INTRODUCTION

Many results in the rapidly evolving field of the neuroscience of music demonstrate that musical practice has important consequences on the anatomo-functional organization of the brain. From an anatomical perspective, magnetic resonance imaging, for instance, has revealed morphological differences between musicians and nonmusicians in auditory (including Heschl's gyrus and secondary auditory cortex), motor (central), and visuospatial (parietal) brain areas (Gaser et al., 2003; Schneider et al., 2002), as well as in the size of the corpus callosum and planum temporale (Schlaug, Jancke, Huang, & Steinmetz, 1995; Schlaug, Jancke, Huang, Staiger, & Steinmetz, 1995). Such anatomical differences have functional implications. Indeed, research using functional magnetic resonance imaging and magnetoencephalography has shown increased activity in Heschl's gyrus of professional and amateur musicians compared with nonmusicians (Schneider et al., 2002), increased somatosensory and motor representations with musical practice (Pantev et al., 1998; Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995), and larger bilateral activation of planum

temporale for musicians than nonmusicians (Ohnishi et al., 2001).

Interestingly, although these different regions may fulfill different musical functions, such as the encoding of auditory information (Heschl's gyrus and secondary auditory cortex), transcoding visual notation into motor representations, and playing an instrument (visuospatial, somatosensory, and motor brain areas), they are not necessarily specific to music. Rather, these different brain structures have also been shown to be activated by other cognitive functions. For example, Heschl's gyrus, the secondary auditory cortex, and planum temporale are typically involved in different aspects of language processing (Meyer, Alter, Angela, Lohmann, & von Cramon, 2002; Tzourio et al., 1997). In addition, visuospatial areas in the parietal lobes have been shown to be activated by approximate calculation in arithmetic (Culham & Kanwisher, 2001; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). Conversely, recent results obtained with the magnetoencephalography method have demonstrated that Broca's area is not as language-specific as believed for almost a century. Indeed, this brain area was activated not only by syntactic processing of linguistic phrases, but also by syntactic processing of musical phrases (Maess, Koelsch, Gunter, & Friederici, 2001).

¹Institut de Neurosciences Cognitives de la Méditerranée,

²Université de la Méditerranée

Taken together, these results show that musical practice has consequences on the anatomo-functional organization of brain regions that are not necessarily specific to music. The idea that we wanted to test in the present experiment is that musical practice, by favoring the development and functional efficiency of specific brain regions, may not only benefit different aspects of music processing, but may also favor positive transfers in other domains of cognition.

Positive transfer due to extended musical practice has been described at the behavioral level, in both adults and children, in domains that are not directly linked to music, such as mathematical abilities (Bilhartz, Bruhn, & Olson, 2000; Costa-Giomi, 1999; Graziano et al., 1999; Gardiner, Fox, Knowles, & Jeffrey, 1996), mental imagery (Aleman, Nieuwenstein, Böcker, & Hann, 2000), symbolic and spatio-temporal reasoning (Gromko & Poorman, 1998; Rauscher et al., 1997), visuo-spatial abilities (Brochard, Dufour, & Desprès, 2004; Cupchick, Philips, & Hill, 2001; Hetland, 2000), verbal memory (Ho, Cheung, & Chan, 2004; Chan, Ho, & Cheung, 1998), self-esteem (Costa-Giomi, 2004), and very recently for measures of general intelligence (Schellenberg, 2004). However, as noted by Thompson, Schellenberg, and Husain (2004), although most of the studies reported above were successful in showing positive correlations between music and other cognitive domains, very few studies have aimed at testing specific hypotheses regarding the causal links underlying these effects. Clearly, such causal links would be easier to test by studying positive transfer between music and other cognitive domains that involve, at least partially, a similar set of computations. One such candidate is language. Indeed, several authors have emphasized the similarities between language and music processing (see Koelsch, 2005; Patel, 2003a, 2003b; Zatorre et al., 2002; Besson & Schön, 2001, for reviews).

Although a number of experiments have aimed at comparing aspects of music and language processing that are presumably quite different, such as syntax and harmony or semantic and melody (Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Besson & Faïta, 1995), only few recent studies have compared two aspects that are objectively more similar, melody and prosody, the music of speech. Prosody has both a linguistic and an emotional function and can broadly be defined at the abstract, phonological level, as the patterns of stress and intonation in a spoken language, and at the concrete, acoustic level, by the same parameters that define melody (i.e., the rhythmic succession of pitches in music), that is, fundamental frequency (F0), intensity, duration, and spectral characteristics. Based on these similarities, Thompson et al. (2004) tackled the emotional function of prosody. They were able to show that adult musicians outperformed adult nonmusicians at identifying emotions (e.g., sadness, fear) conveyed by spoken sentences and by tone sequences that mimicked

the utterances' prosody. Most importantly, they also showed that 6-year-olds, tested after a year of musical training, were better than nonmusician children at identifying anger or fear.

Analyzing both the behavioral measures and variations in brain electrical activity time-locked to events of interest (i.e., event-related brain potentials, or ERPs), Schön, Magne, and Besson (2004) designed an experiment to directly compare pitch processing in music and language (F0). Short musical and linguistic phrases were aurally presented, and the final word/note was melodically/prosodically congruous or incongruous. Incongruities were built by increasing the pitch of the final notes or the F0 of the final words by one fifth of a tone and 35%, respectively, for the weak incongruities and by half of a tone and 120%, respectively, for the strong incongruities. The general hypothesis is that if similar processes underlie the perception of pitch in language and music, then improved pitch perception in music, due to musical expertise, may extend to pitch perception in language. Consequently, musicians should perceive pitch deviations better than nonmusicians not only in music, but also in language. Indeed, results showed that adult musicians not only detected variations of pitch in melodic phrases better than nonmusicians, but that they also detected variations of fundamental frequency in sentences (linguistic prosody) better than nonmusicians. Moreover, detailed analysis of the ERPs revealed that the latency of the positive components elicited by the weak and strong incongruities in both music and language was shorter for musicians than for nonmusicians. Finally, analysis of the amplitude and scalp distribution of early negative components also revealed evidence for positive transfer between music and language.

Based on these results, the aim of the present experiment is twofold. First, we wanted to determine whether such positive transfer effects between pitch processing in music and language would also be found in 8-year-old children. In other words, would 3 to 4 years of extended musical practice be sufficient for musician children to outperform nonmusician children in the detection of pitch violations in both music and language, as was shown for adults with an average of 15 years of musical training (Schön et al., 2004)? Based on the provocative results by Thompson et al. (2004), demonstrating that 1 year of musical training has a strong influence on the identification of emotional prosody, we also expected to find positive evidence for linguistic prosody. Moreover, by using a parametric manipulation of pitch in both language and music as in our previous study (Schön et al., 2004), we were able to make specific predictions regarding the effects of musical training. Thus, we expected no differences between musician and nonmusician children in the detection of congruous endings, because they match the expectations derived from the previous linguistic or musical contexts. Similarly, we expected no differences between the two groups in

the detection of the strong incongruity, because in both language and music, this deviation was constructed in such a way as to be obvious. By contrast, we expected differences between musicians and nonmusician children in the detection of the weak incongruity because this deviation was subtle and should require a musical ear to be detected.

The second aim was to study the neurophysiological basis of positive transfer using a developmental approach. Indeed, one further reason to test 8-year-olds is that previous results, based on the analysis of the auditory evoked potentials, have shown that the auditory cortex is not completely mature at this age (Pang & Taylor, 2000; Ponton, Eggermont, Kwong, & Don, 2000). Typically, the amplitude of the P1, N1b, and P2 components of the auditory evoked potential increases until the age of 10–12 years and remains stable (N1b and P2) or decreases (P1) during adulthood. Moreover, while P1 and N1 latencies typically decrease, P2 latency remains stable and N2 latency increases as a function of age. Thus, it was of interest to compare the ERP effects found in children during the critical period of development of the auditory cortex with those previously found in adults.

RESULTS

Behavioral Data

Results of a three-way analysis of variance (ANOVA) [expertise (two levels), material (two levels), and congruity (three levels)] on the transformed percentages of error showed main effects of expertise [$F(1,18) = 16.59, p < .001$], material [$F(1,18) = 30.53, p < .001$], and congruity [$F(2,36) = 36.05, p < .001$]. Clearly, nonmusician children (27%) made overall more errors than musician children (12%), and both made more errors with the musical (27%) than linguistic materials (12%). Moreover, the error rate was highest for the weak incongruity (see Figure 1). Most importantly, and as predicted, musician children detected the weak incongruity better than nonmusician children, not only in

music, but in language as well [Expertise \times Congruity interaction: $F(2,36) = 4.47, p = .01$, with no Expertise \times Material \times Congruity interaction, $p < .38$].

Electrophysiological Data

Mean amplitude ERPs to final note/word were measured in several latency bands (100–200, 200–400, and 400–700 msec) determined from both visual inspection and based on previous results. Results were analyzed separately for musicians and nonmusicians and for the linguistic and musical materials,¹ using ANOVAs that included congruity (three levels: congruous, weakly incongruous, and strongly incongruous) and electrodes (four levels: Fz, Cz, Pz, and Oz) as within-subject factors for midline analyses. ANOVAs were also computed for lateral electrodes, using six regions of interest (ROIs): left and right fronto-central (F3, F7, Fc5, and F4, F8, Fc6, respectively), left and right temporal (C3, T3, Cp5, and C4, T4, Cp6, respectively), and left and right temporo-parietal (Cp1, P3, T5, and Cp2, P4, T6, respectively). ANOVAs were computed for lateral electrodes using congruity (three levels), hemispheres (two levels: left and right), localization (three levels: fronto-central, temporal, and temporo-parietal), and electrodes (three for each ROI, as described above) for lateral analyses. All p values were adjusted with the Greenhouse–Geisser epsilon correction for nonsphericity when necessary. When the factor congruity was significant or interacted with other factors, planned comparisons between pairs of conditions were computed. To simplify the presentation of the results, outcomes of the main ANOVAs in the different latency ranges are reported in Tables 1 and 2. When the main effects or interactions are significant, results of two by two comparisons are presented in text.

Music

For musician children, the ERPs associated to the final notes clearly differ as a function of congruity in all latency bands (100–200, 200–400, and 400–700 msec)

Figure 1. Percentage of error rates for congruous (Cong) final notes or words and for weak and strong incongruities in music and language are presented for musicians and nonmusicians. Clearly, in both music and language, the percentage of errors to weak incongruities was significantly higher for nonmusicians than for musicians.

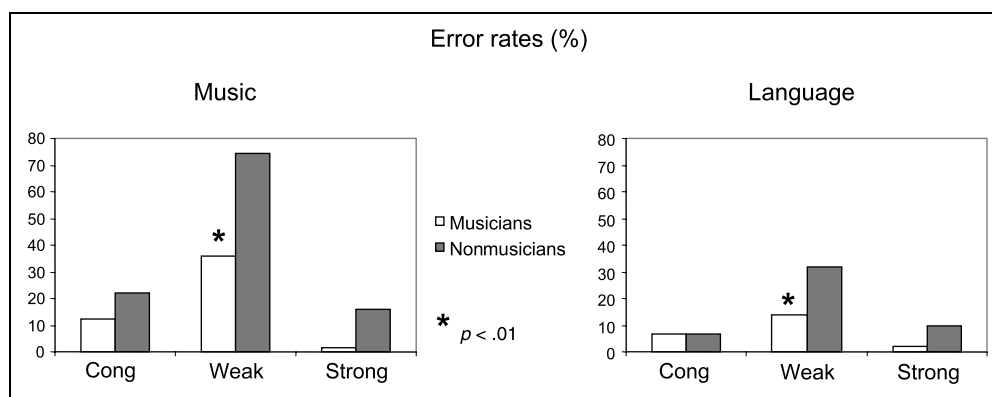


Table 1. Results of Main ANOVAs for Music

| Latency Bands | Electrodes | Factors | Musicians | Nonmusicians |
|---------------|------------|---------|-----------------------------|----------------------------|
| 100–200 msec | Midlines | C | $F(2,18) = 10.44, p = .001$ | <i>ns</i> |
| | Laterals | C | $F(2,18) = 11.83, p < .001$ | <i>ns</i> |
| 200–400 msec | Midlines | C | $F(2,18) = 16.50, p < .001$ | <i>ns</i> |
| | Laterals | C | $F(2,18) = 5.12, p = .019$ | <i>ns</i> |
| 400–700 msec | Midlines | C | $F(2,18) = 10.93, p = .001$ | $F(2,18) = 4.59, p = .026$ |
| | Laterals | C × L | $F(4,36) = 4.50, p = .010$ | $F(4,36) = 5.13, p = .014$ |

C = Congruity, L = localization (three regions of interest: fronto-central, temporal, and temporo-parietal).

considered for analysis (see Table 1 for results of main ANOVAs). Compared with congruous notes, weak incongruities elicited a larger early negative component, between 200 and 400 msec, with maximum amplitude around 340 msec [midlines: $F(1,9) = 23.20, p < .001$; laterals: $F(1,9) = 6.92, p < .027$; see Figures 2 and 3]. This negative effect was well distributed over the scalp, as suggested by the absence of any significant Congruity × Localization interactions at lateral electrodes.

Strong incongruities also elicited a larger early negative component than congruous notes, with maximum amplitude around 210 msec. This effect was significant earlier, between 100 and 200 msec, than for the weak incongruities and was broadly distributed across scalp sites [midlines: $F(1,9) = 22.56, p < .001$; laterals: $F(1,9) = 19.23, p < .001$; no Congruity × Localization interaction at lateral electrodes; see Figure 2]. Moreover, this early negative component was followed by an increased positivity that differed from the ERPs to congruous notes as early as 200–400 msec at midline sites [$F(1,9) = 6.06, p = .036$]. This effect extended in the 400- to 700-msec range and was significant at both midlines [$F(1,9) = 19.67, p < .001$] and lateral electrodes [Congruity × Localization interaction: $F(2,18) = 8.01, p = .003$], with a temporo-parietal distribution [$F(1,9) = 11.89, p = .007$; see Figure 3].

In contrast to musician children, the ERPs to weak incongruities in nonmusicians did not differ from congruous notes in any of the latency bands considered for analysis (see Figure 2). However, strong incongruities elicited an early negative component, peaking around 250 msec. This effect was significant later (in the 200- to 400-msec latency band) than in musicians and was larger over the right hemisphere [Congruity × Hemisphere interaction: $F(1,9) = 4.97, p = .05$, see Figures 2 and 3]. This early negative component was also followed by an increased positivity compared with congruous note, but that started later, in the 400- to 700-msec range, than for musician children [midlines: $F(1,9) = 7.82, p = .02$; laterals: Congruity × Localization interaction, $F(2,18) = 13.71, p = .003$]. This positive effect was localized over the temporal and temporo-parietal sites bilaterally [temporal: $F(1,9) = 17.77, p = .002$; temporo-parietal: $F(1,9) = 22.24, p = .001$, see Figures 2 and 3].

Language

For musician children, the main effect of congruity was significant in both the 200- to 400-msec and the 400- to 700-msec ranges (see Table 2). Both weak and strong prosodic incongruities elicited larger positivities than congruous endings between 200 and 700 msec (see

Table 2. Results of Main ANOVAs for Language

| Latency Bands | Electrodes | Factors | Musicians | Nonmusicians |
|---------------|------------|---------|-----------------------------|----------------------------|
| 100–200 msec | Midlines | C | <i>ns</i> | <i>ns</i> |
| | Laterals | C | <i>ns</i> | <i>ns</i> |
| 200–400 msec | Midlines | C | $F(2,18) = 10.88, p = .001$ | $F(2,18) = 3.86, p = .041$ |
| | Laterals | C × L | $F(4,36) = 4.65, p = .022$ | <i>ns</i> |
| 400–700 msec | Midlines | C | $F(2,18) = 11.84, p < .001$ | $F(2,18) = 4.27, p = .032$ |
| | Laterals | C | $F(2,18) = 5.67, p = .015$ | <i>ns</i> |
| | | C × L | $F(4,36) = 6.69, p = .006$ | $F(4,36) = 3.90, p = .043$ |

C = Congruity, L = localization (three regions of interest: fronto-central, temporal, and temporo-parietal).

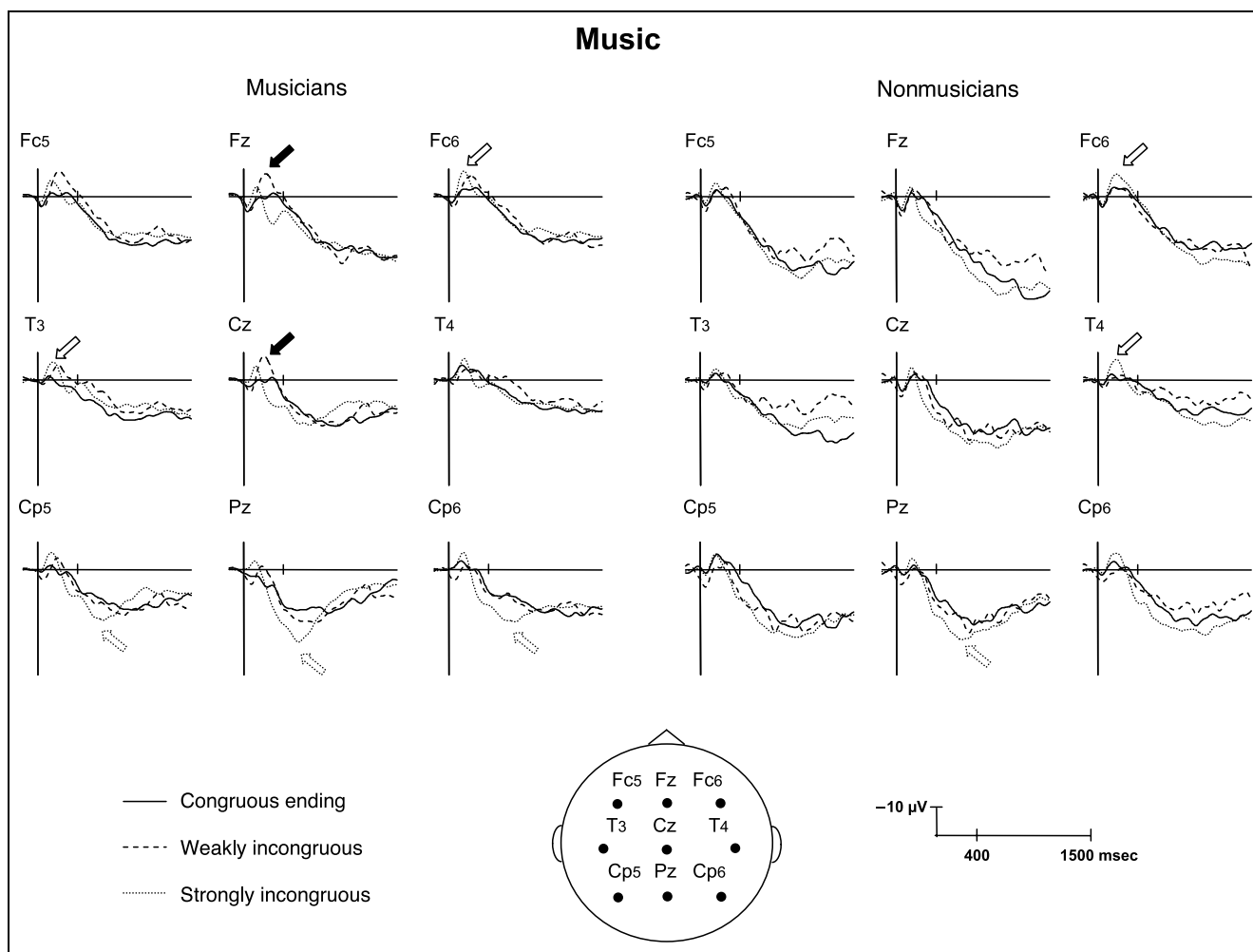


Figure 2. Illustration of the variations in brain electrical activity time-locked to final note onset and elicited by congruous endings, weak incongruities, or strong incongruities. Each trace represents an average of electrophysiological data recorded from 10 musician and 10 nonmusician 8-year-old children. EEG was recorded from 28 electrodes; selected traces from 9 electrodes are presented. In this figure, as in the following ones, the amplitude (in microvolts) is plotted on the ordinate (negative up) and the time (in milliseconds) is on the abscissa. White arrows point to the effects that are present for both musician and nonmusician children, whereas black arrows show effects that are present for musicians only.

Figure 4). This positive effect was largest over the midline sites for the weak incongruity [200–400 msec: $F(1,9) = 9.42, p = .013$; 400–700 msec: $F(1,9) = 8.90, p = .015$; see Figure 5] and was broadly distributed over the scalp, with a bilateral temporo-parietal distribution for the strong incongruities (Congruity \times Localization interaction, 200–400 msec: $F(2,18) = 8.18, p = .015$, 400–700 msec: $F(2,18) = 16.65, p < .001$; results of post hoc comparisons in the temporal and temporo-parietal ROIs always revealed significant differences at $p < .05$).

Although the main effect of congruity was also significant for nonmusician children in both the 200- to 400-msec and the 400- to 700-msec ranges (see Table 2) results of 2×2 comparisons showed that only the ERPs associated to strong incongruities elicited larger positivities than congruous endings (see Figure 4). This effect was significant between 200 and 700 msec at midline sites [200–400 msec: $F(1,9) = 9.57, p = .012$; 400–700 msec:

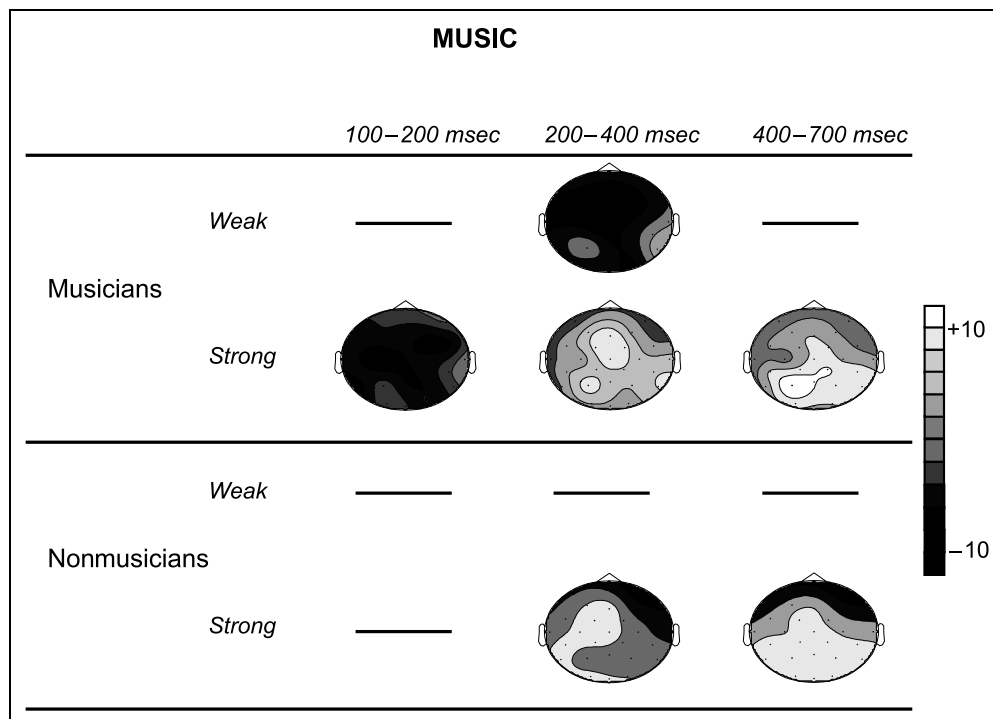
$F(1,9) = 11.27, p = .008$] and between 400 and 700 msec at lateral sites [Congruity \times Localization interaction: $F(2,18) = 12.32, p < .001$], with a bilateral temporo-parietal maximum [$F(1,9) = 15.56, p = .003$; see Figure 5].

Finally, results of ANOVAs performed in successive 25-msec latency bands between 200 and 400 msec over the midline sites revealed that the positive differences between strong incongruities and congruous endings started earlier for musician (275–300 msec, $p < .01$) than nonmusician children (350–375 msec, $p < .01$; see Table 3).

DISCUSSION

In line with our hypotheses, error rate analyses showed that musician children outperformed nonmusician children in the detection of weak incongruities, not only in music, but also in language, thereby pointing to a

Figure 3. Topographic maps of the weak incongruity effect (mean amplitude difference between weak incongruity and congruous ending) and strong incongruity effect (mean amplitude difference between strong incongruity and congruous ending) in music for musicians (top) and nonmusicians (bottom). In the three latency windows considered for analyses (100–200, 200–400, and 400–700 msec), only significant effects are represented.



common pitch processing mechanism in language and music perception. In line with these behavioral data, ERPs analyses also showed greatest differences between the two groups of children for the weak incongruity. In this case, both an early negative component in music and a late positive component in language were found only for musician children. By contrast, early negative and late positive components were elicited by strong incongruities in both groups, although with some quantitative differences. These results are considered in turn in the following discussion.

Effects of Musical Training on the Detection of Pitch Changes in Music and Language

Behavioral data clearly showed that the overall level of performance in the pitch detection task was higher for musician than nonmusician children. This difference was expected in the music task because musician children had 4 years of musical training on average, and previous reports have highlighted the positive effect of musical expertise on music perception in both adults and children (Schön et al., 2004; Thompson, Schellenberg, & Husain, 2003, 2004; Besson & Faïta, 1995, but see also Bigand, Parncutt, & Lerdahl, 1996, for evidence in adults that suggests otherwise). What is most striking is that musicians' performance was also better in language. Because Schellenberg (2004) recently showed that 1 year of musical training significantly improved IQ, one could argue that general nonspecific processes are at play, which explains why musician children outperformed

nonmusician children. In this case, however, one would expect differences between the two groups of children in the three experimental conditions. The present results show that this is not the case: The only significant difference between the two groups was found for the weak incongruity, which is clearly the most difficult to detect. In this condition, for both music and language, the level of performance of musician children was twice as high as for nonmusician children. Therefore, although music training may improve general intelligence (Schellenberg, 2004), it also seems to exert specific beneficial influences on both music and language perception. Although positive transfer effects between music and other cognitive domains have already been reported in the literature, as mentioned in the Introduction, the causal links underlying these effects have not been directly tested. Here, we provide evidence that music training, by increasing sensitivity to a specific basic acoustic parameter, pitch, which is equally important for music and speech prosody, does enhance children's ability to detect pitch changes not only in music, but also in language.

The present results also extend those recently reported by Thompson et al. (2004), which showed that 1 year of musical training allowed 6-year-olds to identify emotional prosody in utterances better than the nonmusician control group. Thus, evidence for positive transfer effects between music and language is increasingly being shown when basic acoustic parameters, such as pitch, intensity, or duration, are manipulated in both domains.

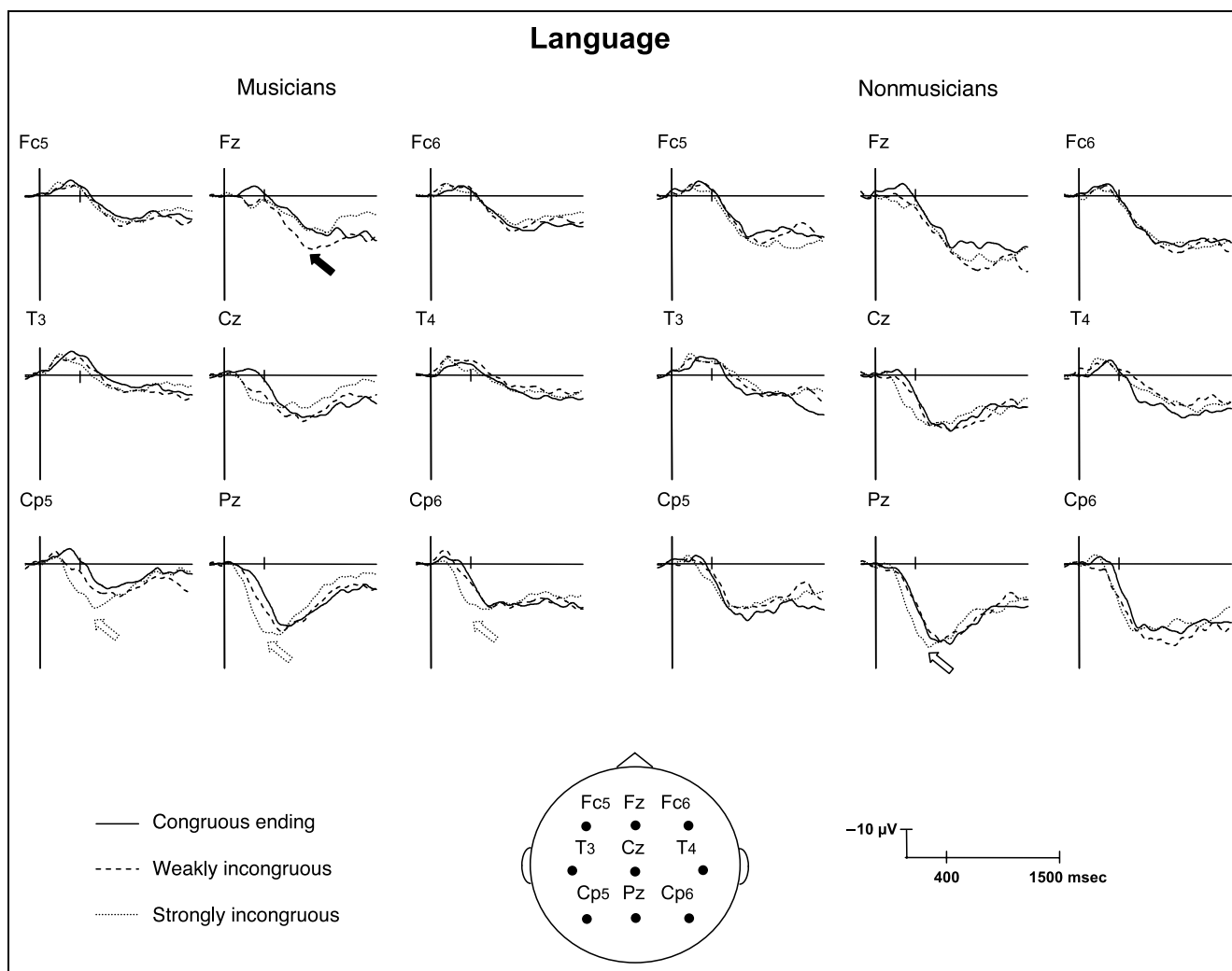


Figure 4. Illustration of the variations in brain electrical activity time-locked to final word onset and elicited by congruous endings, weak incongruities, or strong incongruities. Each trace represents an average of electrophysiological data recorded from 10 musician and 10 nonmusician children.

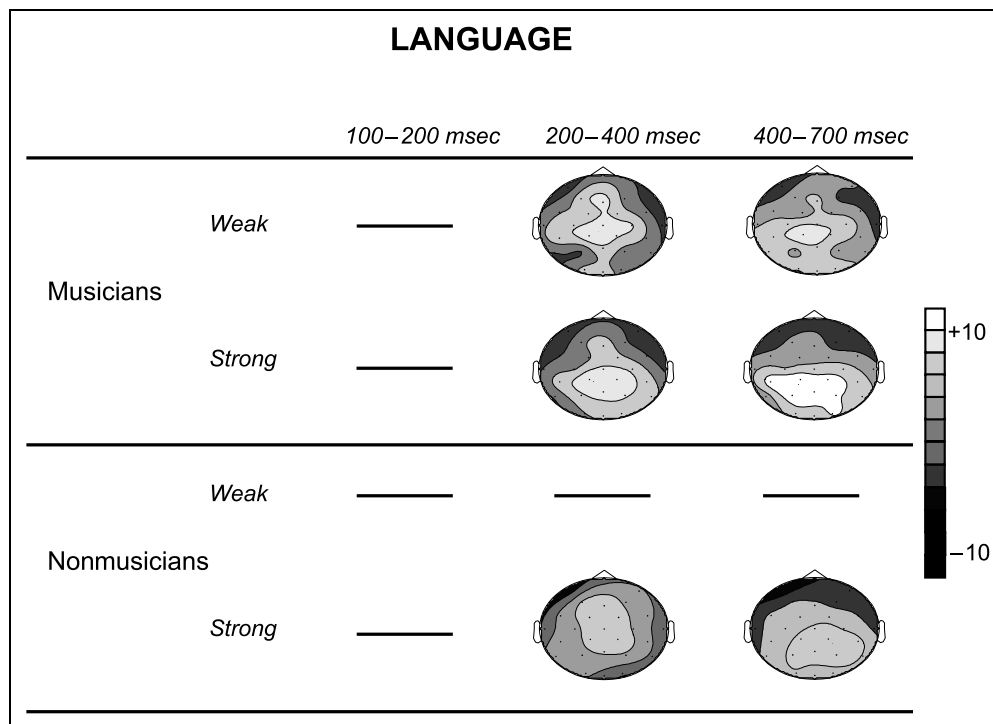
Neurophysiological Basis of Positive Transfer Effects

In line with the behavioral data, ERPs analyses showed that the differences between musician and nonmusician children were larger for the weak incongruity than for both congruous endings and strong incongruities. Indeed, although for musician children, weak incongruities elicited a larger early negativity than congruous notes in music and a larger late positivity than congruous words in language, no such differences were found for nonmusician children in either music or language (see Figures 3 and 5). Therefore, when pitch violations are most difficult to detect, different processes seem to be involved as a function of musical expertise.

By contrast, similar ERP patterns were elicited by the strong incongruity for both musician and nonmusician children. In music, early negativities were followed by late positivities in both groups. Importantly, however,

precise analyses of their time course and scalp distribution also revealed some quantitative differences. First, the onset of the early negative effect was 100 msec shorter (significant between 100 and 200 msec for musicians and between 200 and 400 msec for nonmusicians) and the onset of the late positive effect was 200 msec shorter (significant between 200 and 400 msec for musicians and between 400 and 700 msec for nonmusicians) for musician than nonmusician children. Second, although the early negative effect was broadly distributed over the scalp for musician children, it was localized over the right hemisphere for nonmusician children. Although right lateralization for pitch processing is in line with some results in the literature (see Zatorre et al., 2002), the lateralized distribution reported here may result from an overlap of the early negative components by subsequent later positivities that developed over left fronto-central regions, thereby reducing the negativity over the left hemisphere (see Figure 3).

Figure 5. Topographic maps of the weak incongruity effect (mean amplitude difference between weak incongruity and congruous ending) and strong incongruity effect (mean amplitude difference between strong incongruity and congruous ending) in language for musicians (top) and nonmusicians (bottom). Only significant effects are represented.



In language, the strong incongruity elicited a larger late positive component than the congruous word for both musician and nonmusician children. The precise analysis of the time course of this positive effect revealed that it started 75 msec earlier and was larger for the musician than nonmusician groups, although with a similar temporo-parietal scalp distribution.

Taken together, these results clearly show both qualitative (scalp distribution) and quantitative (latency differences) differences between musician and nonmusician children. It is also interesting to note that the late positivity is overall larger and lasts longer for nonmusicians than musicians, which may reflect the fact that nonmusicians need more processing resources to per-

form the tasks and that processing takes longer than for musicians.

Developmental Perspective

The functional significance of the results reported above is now considered in light of previous results found with adults performing the same tasks with the same materials (Schön et al., 2004). Considering first the differences, the overall ERP amplitude was larger and the latency of the ERP components was longer in children than in adults. Consider, for instance, the negative component to strong musical incongruity at the electrode T4 where it is clearly defined. The mean amplitude of this negativity was $-7.35 \mu\text{V}$, and its peak latency 255 msec for children (musicians and nonmusicians), whereas it was $-4.10 \mu\text{V}$ and 165 msec, for adults. These results are in line with a large literature showing decreased amplitude and shortened latency of ERP components as age progresses (see Taylor, 1995, for a review). Decreases in amplitude are thought to depend upon the number of pyramidal cell synapses contributing to postsynaptic potentials (Ponton et al., 2000) and are interpreted as reflecting the automation of the underlying processes that thereby require fewer and fewer neurons (Batty & Taylor, 2002). Decreases in latency may result from increased speed of nervous transmission, due to axon myelination, as well as to the maturation of synaptic connections, due to the repeated synchronization of specific neuronal populations (Batty & Itier, 2004; Taylor, 1995; Courchesne, 1990; Eggermont, 1988). In sum, the overall decreased amplitude and shortened

Table 3. Timing of the Strong Incongruity Effect in Language

| Latencies (msec) | Musicians | Nonmusicians |
|------------------|-----------|--------------|
| 200–225 | – | – |
| 225–250 | – | – |
| 250–275 | – | – |
| 275–300 | * | – |
| 300–325 | * | – |
| 325–350 | ** | – |
| 350–375 | ** | * |
| 375–400 | ** | ** |

* $p < .01$.

** $p < .001$.

latency of ERPs with age may reflect an enhanced efficiency of cognitive processing over the course of development. Moreover, the differences between musicians and nonmusician children in the amplitude and latency of the early negative components elicited by the weak and strong incongruities in music are in line with recent results by Shahin, Roberts, and Trainor (2004) showing overall enhanced amplitude of the early ERP components with musical practice in 4- to 5-year-old children. Interestingly, these results also showed that the increase in amplitude of the N1 and P2 components was specific to the instrument played.

Regarding the present series of experiments, results revealed differences in the early negative components between the adults tested by Schön et al. (2004) and the children tested here when they perform the same explicit task (pitch congruity judgment) on the same materials. In adults, early negative components were elicited, between 50 and 200 msec, by strong incongruities both in music and in language. In music, they were distributed over the right temporal regions, whereas they were distributed over the temporal regions bilaterally in language. By contrast, for children, they were only found in music.

The functional interpretation of these early negativities is still a matter of debate. Previous results in adults have shown that both harmonic (Koelsch, Gunter, Friederici, & Schröger, 2000; Patel et al., 1998) and melodic incongruities (Schön et al., 2004) elicit an early negative component over right frontal sites between 200 and 400 msec. Moreover, results of a study with 5- and 9-year-old nonmusician children have also shown that harmonic violations elicited early negative components (Koelsch et al., 2003). Finally, the finding that these early negativities were typically elicited in musical contexts and were larger for participants with than without formal musical training led the authors to propose that this early negativity may reflect specific musical expectancies (Koelsch, Schmidt, & Kansok, 2002). However, many results in the literature have also demonstrated that unexpected changes in the basic acoustic properties of sounds, such as frequency, intensity, or duration, elicit an early automatic brain response, the mismatch negativity (Näätänen, 1992). Therefore, the issue of whether the early negativity reflects specific musical expectancies or a domain general mismatch detection process remains an open question (Koelsch, Maess, Grossmann, & Friederici, 2002; Koelsch, Schröger, & Gunter, 2002). Because early negative components were found in response to pitch violations in both language and music in our previous experiment with adults (Schön et al., 2004), we favor the interpretation following which they reflect automatic aspects of pitch processing in both domains. However, how can we reconcile such an interpretation with the present results with children showing an early negativity to pitch deviations in music, but no such component in language? This

matter raises the intriguing possibility that automatic detection of pitch changes in music may be functional earlier on (as early as 5–8 years old) than in language. Several authors have emphasized the importance of melodic elements in infant-directed speech and for language acquisition (Trehub, 2003; Papousek, 1996; Jusczyk & Krumhansl, 1993). Thus, the development of the early negativity in both music and language needs to be tested in further experiments using a longitudinal approach with children ages 4, 6, 8, and 10 years.

Turning to the similarities, results with children showed that, as was previously found with adults, strong incongruities elicited late positive components with a centro-parietal distribution in both music and language. Therefore, in contrast with the processes underlying the negative components, the processes underlying the occurrence of these late positivities seem to be present already at age 8 years in both music and language. Based on numerous results in the ERP literature, the occurrence of these late positivities (P3b component) is generally considered as being related to the processing of surprising and task-relevant events (Picton, 1992; Duncan-Johnson & Donchin, 1977; see Donchin & Coles, 1988, for a review). Moreover, the latency of these positive components often varies with the difficulty of the categorization task (Kutas, MacCarthy, & Donchin, 1977), which is in line with our previous and present results showing shorter latencies for the strong than weak incongruities.

Conclusions

The most important conclusion to be drawn from these results is that we found behavioral evidence for a common pitch processing mechanism in language and music perception. Moreover, by showing qualitative and quantitative differences in the ERPs recorded from musician and nonmusician children, we were able to uncover some of the neurophysiological processes that may underlie positive transfer effects between music and language. The occurrence of an early negative component to the weak incongruity in music for musician children only may indeed reflect a greater sensitivity to pitch processing. Such enhanced pitch sensitivity in musician children would also be reflected by the larger late positivity to weak incongruities than congruous words in language that was not found in nonmusician children. Although these findings may reflect the facilitation, due to musical training, of domain-general pitch mismatch detection processes common to both music and language, further experiments are needed to specify the relationships between the early negative and late positive components and why early negative components were elicited by strong incongruities in both musician and nonmusician children in music but not in language.

To summarize, these results add to the body of cognitive neuroscience literature on the beneficial effects of musical education; in particular, the present findings highlight the positive effects of music lessons for linguistic abilities in children. Therefore, these findings argue in favor of music classes being an intrinsic and important part of the educational programs in public schools and in all the institutions that aim at improving children's perceptive and cognitive abilities. Finally, the present study also confirms that the ERP method is particularly well adapted for the exploration of positive transfer effects between music processing and other cognitive domains. Further research is also needed to determine the extent of these transfers, as well as their existence between music cognition and nonauditory processes such as visuospatial reasoning.

METHODS

Participants

Twenty-six children (14 girls and 12 boys; age 8 ± 1 years), 13 musicians and 13 nonmusicians, participated in the experiment, which lasted for about 2 hr. The musician children had 4 ± 1 years of musical training on average. All children were right-handed, had normal hearing, and were native speakers of French. Most importantly, all the children came from the same elementary school and had similar socioeconomic backgrounds (e.g., a *t* test on the mean family incomes revealed no significant differences between the two groups, $p = .74$). All musician children played an instrument (violin = 5, guitar = 2, flute = 1, clarinet = 2, harp = 1, piano = 2), which they regularly practiced

everyday for around 20 to 30 min. They also took music lessons twice a week for a half an hour. Thus, these children played music for about 3–4 hr per week. All nonmusician children also had regular extracurricular activities (judo = 2, swimming = 2, cycling = 2, tennis = 1, rugby = 1, rollerblading = 1, circus training = 1, gymnastics = 1, horseback riding = 1, soccer = 1). Six of the participants (three musicians and three nonmusicians) were not included in the analyses because of technical problems or too many artifacts during the electroencephalogram (EEG) recording session. Children were given presents at the end of the recording session. All parents gave informed consent for their children to participate in the experiment.

Stimuli

Stimuli comprised 96 French-spoken declarative sentences taken from children's books and ending with bisyllabic words (e.g., “Dans la barque se tient l'ennemi de Peter Pan, le terrible pirate”/“In the boat is the enemy of Peter Pan, the terrible pirate”). Sentences were spoken at a normal speech rate by a native French female speaker, recorded in a soundproof room using a digital audiotape (sampling at 44.1 kHz), and synthesized using the software Winpitch (Martin, 1996). The mean duration of the sentence was 3.97 ± 0.7 sec.

A total of 96 melodies were also presented in the experiment. Half were selected from the repertoire of children's music (e.g., “Happy Birthday”), and half were composed for the experiment by a professional musician, following the same rules of composition as for familiar melodies. Tunes were converted into MIDI files



Figure 6. Examples of stimuli used in the experiment. (A) The speech signal is illustrated for the sentence: “Un loup solitaire se faufile entre les troncs de la grande forêt” [literal translation: “A lonely wolf worked his way through the trees of the big forest”]. (B) The musical notation is illustrated for the song “Happy Birthday.”

using the synthetic sound of a piano (KORG XDR5, Tokyo, Japan). The mean duration of the melodies was 10.3 ± 2.44 sec.

An equal number of sentences/melodies (32) were presented in each of the three following experimental conditions, thus leading to a total of 192 stimuli with 96 sentences and 96 melodies: The final word or note was prosodically or melodically congruous, weakly incongruous, or strongly incongruous (see Figure 6A). Based upon results of pretests of a preliminary version of this material with both adults and children, the F0 of the last word was increased, using the software WinPitch, by 35% for the weak incongruity and by 120% for the strong incongruity (without changing the original pitch contour). In the musical material, the last note was increased by one fifth of a tone for the weak incongruity and by half of a tone for the strong incongruities using the sound file editor software Wavelab, Hamburg, Germany (see Figure 6B).

Procedure

In eight separate blocks of trials, children were required to listen attentively, through headphones, either to the melodies (four blocks) or the sentences (four blocks). Within each block of trials, stimuli were presented in a pseudorandom order, and children were asked to decide whether the last word or note seemed normal or strange (i.e., something was wrong), by pressing one of two response keys as quickly and as accurately as possible. The hand of response and the order of presentation (musical or prosodic materials first) were counterbalanced across children.

Event-related Brain Potential Recordings

EEG was recorded for 2200 msec starting 150 msec before the onset of the last word/note, from 28 scalp electrodes, mounted on a child-sized elastic cap and located according to the International 10/20 system. These recording sites plus an electrode placed on the right mastoid were referenced to the left mastoid electrode. The data were then rereferenced offline to the algebraic average of the left and right mastoids. Impedances of the electrodes never exceeded 3 k Ω . To detect blinks and vertical eye movements, the horizontal electrooculogram (EOG) was recorded from electrodes placed 1 cm to the left and right of the external canthi, and the vertical EOG was recorded from an electrode beneath the right eye, referenced to the left mastoid. Trials containing ocular or movement artifacts, or amplifier saturation, were excluded from the averaged ERP waveforms. The EEG and EOG were amplified by an SA Instrumentation amplifier with a bandpass of 0.01–30 Hz and were digitized at 250 Hz by a PC-compatible microcomputer (Compaq Prosignia 486, Hewlett-Packard Co., Palo Alto, CA).

Acknowledgments

This research was first supported by a grant from the International Foundation for Music Research (IFRM: RA 194) and later by a grant from the Human Frontier Science Program to Mireille Besson (HSFP: RGP0053). Cyrille Magne benefited from a research fellowship from the Cognitive Program of French Ministry of Research, and Daniele Schön was a post-doctorate student supported by the HFSP grant. The authors acknowledge Monique Chiambretto and Reyna Leigh Gordon for their technical assistance.

Reprint requests should be sent to Cyrille Magne, Center for Complex Systems and Brain Sciences, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA, or via e-mail: magne@ccs.fau.edu.

Note

1. Because latency, amplitude, and scalp distribution differences were found between musician and nonmusician children and between the linguistic and musical materials, results of a general ANOVA including expertise and materials as factors are less informative than computing the analyses for each group and each material separately.

REFERENCES

- Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., & Hann, E. H. F. (2000). Music training and mental imagery ability. *Neuropsychologia*, *38*, 1664–1668.
- Batty, B., & Itier, R. J. (2004). Les modifications des potentiels évoqués cognitifs au cours du développement. In B. Renault (Ed.), *L'imagerie fonctionnelle électrique (EEG) et magnétique (MEG): Ses applications en sciences cognitives* (pp. 217–234). Paris: Hermès.
- Batty, M., & Taylor, M. J. (2002). Visual categorization during childhood: An ERP study. *Psychophysiology*, *39*, 1–9.
- Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with non-musicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1278–1296.
- Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*, 232–259.
- Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, *58*, 125–141.
- Bilhartz, T. D., Bruhn, R. A., & Olson, J. E. (2000). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology*, *20*, 615–636.
- Brochard, R., Dufour, A., & Desprès, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain & Cognition*, *54*, 103–109.
- Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature*, *396*, 128.
- Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music*, *32*, 139–152.
- Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, *47*, 198–212.

- Courchesne, E. (1990). Chronology of postnatal human brain development: Event-related potentials, positron emission tomography, myelogenesis, and synaptogenesis studies. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson (Eds.), *Event-related brain potentials* (pp. 210–241). Oxford: Oxford University Press.
- Culham, J. C., & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology*, *11*, 157–163.
- Cupchick, G. C., Phillips, K., & Hill, D. S. (2001). Shared processes in spatial rotation and musical permutation. *Brain & Cognition*, *46*, 373–382.
- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970–974.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context-updating? *Behavioral and Brain Science*, *11*, 355–372.
- Duncan-Johnson, C., & Donchin, E. (1977). On quantifying surprise, the variation of event-related potentials with subjective probability. *Psychophysiology*, *14*, 456–467.
- Eggermont, J. J. (1988). On the maturation of sensory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *70*, 293–305.
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305–307.
- Gardiner, M. F., Fox, A., Knowles, F., & Jeffrey, D. (1996). Learning improved by arts training. *Nature*, *381*, 284.
- Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*, 9240–9245.
- Graziano, A. B., Peterson, M., & Shaw, G. L. (1999). Enhanced learning of proportional math through music training and spatial-temporal training. *Neurological Research*, *21*, 139–152.
- Gromko, J. E., & Poorman, A. (1998). The effect of music training on preschoolers' spatial temporal task performance. *Journal of Research in Music Education*, *46*, 173–181.
- Hetland, L. (2000). Learning to make music enhances spatial reasoning. *Journal of Aesthetic Education*, *34*, 179–238.
- Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*, 439–450.
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 627–640.
- Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*, 207–212.
- Koelsch, S., Grossmann, T., Gunter, T. C., Hahne, A., Schröger, E., & Friederici, A. D. (2003). Children processing music: Electric brain responses reveal musical competence and gender differences. *Journal of Cognitive Neuroscience*, *15*, 683–693.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “Non-musicians” are musical. *Journal of Cognitive Neuroscience*, *12*, 520–541.
- Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. (2002). Sex difference in music-syntactic processing. *NeuroReport*, *14*, 709–712.
- Koelsch, S., Schmidt, B., & Kansok, J. (2002). Influences of musical expertise on the ERAN: An ERP-study. *Psychophysiology*, *39*, 657–663.
- Koelsch, S., Schröger, E., & Gunter, T. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology*, *39*, 1–11.
- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*, 792–795.
- 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.
- Martin, P. (1996). WinPitch: Un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows. *Actes des XXIV Journées d'Etude sur la Parole*, 224–227. Avignon, France.
- Meyer, M., Alter, K., Angela, A. D., Lohmann, G., & von Cramon, D. Y. (2002). fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, *17*, 73–88.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., & Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex*, *11*, 754–760.
- Pang, E. W., & Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: An examination of speech and non-speech stimuli. *Clinical Neurophysiology*, *111*, 388–397.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.
- Papousek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège & J. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 88–112). Oxford, New York.
- Patel, A. D. (2003a). Rhythm in language and music: Parallels and differences. *Annals of the New York Academy of Sciences*, *999*, 140–143.
- Patel, A. D. (2003b). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674–681.
- Patel, A., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717–733.
- Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*, 456–479.
- Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, *111*, 220–236.
- Rauscher, F. H., Shaw, G. L., Levine, L. J., Wright, E. L., Dennis, W. R., & Newcomb, R. (1997). Music training causes long-term enhancement of pre-school children's spatial-temporal reasoning. *Neurological Research*, *19*, 2–8.
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*, 511–514.
- Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*, 699–701.
- Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*, 1047–1055.
- Schneider, P., Scherg, M., Dösch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*, 688–694.

- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Electrophysiological study of pitch perception in language and music. *Psychophysiology*, *41*, 341–349.
- Shahin, A., Roberts, L. E., & Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *NeuroReport*, *15*, 1917–1921.
- Taylor, M. J. (1995). The role of event-related potentials in the study of normal and abnormal cognitive development. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 187–211). Amsterdam: Elsevier.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*, 46–64.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2003). Perceiving prosody in speech: Effects of music lessons. *Annals of the New York Academy of Sciences*, *999*, 530–532.
- Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, *6*, 669–673.
- Tzourio, N., Massiou, F. E., Crivello, F., Joliot, M., Renault, B., & Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *Neuroimage*, *5*, 63–77.
- Zatorre, R. J. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science*, *6*, 37–46.