BRAIN RESEARCH

Research Report

Musical scale properties are automatically processed in the human auditory cortex

Elvira Brattico[a,b,*], Mari Tervaniemi[a,b], Risto Näätänen[a,b], Isabelle Peretz[c]

[a]Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, Finland
[b]Helsinki Brain Research Centre, Finland
[c]Department of Psychology, University of Montreal, Canada

ARTICLE INFO

ABSTRACT

While listening to music, we immediately detect 'wrong' tones that do not match our expectations based on the prior context. This study aimed to determine whether such expectations can occur preattentively, as indexed by event-related potentials (ERPs), and whether these are modulated by attentional processes. To this end, we recorded ERPs in nonmusicians while they were presented with unfamiliar melodies, containing either a pitch deviating from the equal-tempered chromatic scale (out-of-tune) or a pitch deviating from the diatonic scale (out-of-key). ERPs were recorded in a passive experiment in which subjects were distracted from the sounds and in an active experiment in which they were judging how incongruous each melody was. In both the experiments, pitch incongruities elicited an early frontal negativity that was not modulated by attentional focus. This early negativity, closely corresponding to the mismatch negativity (MMN) of the ERPs, was mainly originated in the auditory cortex and occurred in response to both pitch violations but with larger amplitude for the more salient out-of-tune pitch than the less salient out-of-key pitch. Attentional processes leading to the conscious access of musical scale information were indexed by the late parietal positivity (resembling the P600 of the ERPs) elicited in response to both incongruous pitches in the active experiment only. Our results indicate that the relational properties of the musical scale are quickly and automatically extracted by the auditory cortex even before the intervention of focused attention.

1. Introduction

Music is replete with sound events that are cognitively meaningful, creating a vivid internal musical experience in the human mind. In order to deal with the wealth of information impinging on the auditory system, attentive neural mechanisms select and organize the musical input for further cognitive processing by allocating neural resources to the relevant sound events (cf. Coull, 1998).

A central question in cognitive neuroscience concerns the level of attentional control required in input analysis. Behavioral and electrophysiological evidence indicate that several aspects of the auditory environment are analyzed before the intervention of voluntary attention in an automatic and irrepressible way (Velmans, 1991; Näätänen et al., 2001; Schröger et al., 2004). For instance, in the domain of language, dichotic-listening experiments have shown that the meaning of words can be accessed without attention (for a review, see

---

* Corresponding author. Cognitive Brain Research Unit, Department of Psychology, P.O. Box 9 (Siltavuorenpenger 20 C), 00014 University of Helsinki, Finland. Fax: +358 9 191 29450.
E-mail address: elvira.brattico@helsinki.fi (E. Brattico).

Velmans, 1991). What about music? Encoding of both music and language is based on the perceptual analysis of the auditory scene (Bregman, 1990). It also depends on culture-dependent knowledge that is implicitly acquired by exposure (e.g., Krumhansl, 2000; Tillmann et al., 2000; Kuhl, 2004). Likewise with language, the human brain may thus possess neural mechanisms that automatically extract culture-dependent musical information before the intervention of focused attention.

The first step in pitch encoding consists of extracting universal, i.e., culture-independent information, from the music or speech signal. At this stage of analysis, pitch encoding does not require attention. The mismatch negativity (MMN) component of the event-related potential (ERP) (Näätänen et al., 1978; Näätänen and Winkler, 1999) is one of the indices of a mechanism holding and manipulating pitch as well as other sound features in a short time span in order to track change in a repetitive sound environment. Brain evidence based on the MMN component shows that neuronal populations of the auditory cortex react to a simple pitch change in a repetitive sound sequence even outside the focus of attention (e.g., Schröger, 1994; Brattico et al., 2000, 2001; for a review, see Tervaniemi and Brattico, 2004). Such pitch encoding process seems to occur in the neuronal circuits of the primary and secondary areas of the auditory cortex, in particular, in the superior temporal gyrus (Tervaniemi et al., 2000; Müller et al., 2002). The pitch contour of short tone patterns, i.e., changes in pitch direction regardless of pitch distance, are also automatically encoded by the neural circuits underlying MMN elicitation (Tervaniemi et al., 2001; Trainor et al., 2002). Furthermore, pitch relations, that is, musical intervals, are automatically maintained as neural traces even irrespectively of the absolute pitch level of the interval components (Paavilainen et al., 1999; Trainor et al., 2002). These data suggest that, beyond the encoding of absolute pitch by means of the tonotopic maps in the auditory cortex, the central auditory system encodes in a fast and automatic way also the relative distances between pitches (for the role of the frontal cortex in this context, see Korzyukov et al., 2003). Music processing, particularly in its melodic dimension, takes advantage of these automatic auditory mechanisms. For instance, we can immediately recognize a melody despite its transposition to other pitch levels. The presence of monodic (unaccompanied) music in all known cultures as well as the infant's abilities in recognition and memorization of simple melodies (Trehub et al., 1990; Trainor and Trehub, 1993) provides strong support for the elementary nature of the processes involved with melody perception.

In each music culture, however, different sets of pitches are used. The Western tonal system is based on 12 tones fixed according to the equal-tempered tuning, with intervals not smaller than a semitone (also known as half step). In the equal-temperament tuning system, a semitone corresponds exactly to one twelfth of an octave (corresponding to about 6% frequency difference between two tones). The chromatic scale of Western music includes all the 12 tones of the equal tempered tuning system. From this pool of tones, a subset of 7 tones, also termed key or diatonic scale, are usually played in a short piece of music. The tones of the diatonic scale are said to be in-key, whereas the remaining 5 tones are out-of-key. So far, no study has clarified whether the more culture-dependent knowledge of musical scales can be accessed and used to process incoming pitches at a preattentive level. Electrophysiological brain responses to musical scale violations have only been obtained under active paradigms in which subjects were required to judge the congruousness of the sound ending a familiar or unfamiliar melody (Besson et al., 1994; Besson and Faïta, 1995; Hantz et al., 1997). These studies have shown that the out-of-key pitches generate a late, long-lasting positive ERP deflection, termed the P600, peaking at 500–700 ms from sound onset, as compared to in-key pitches (Besson et al., 1994; Besson and Faïta, 1995; Hantz et al., 1997). These paradigms can be regarded as remote from everyday listening situations, however (cf. Schmithorst, 2005).

Pitches are also hierarchically structured according to the rules of tonality and harmony (Krumhansl, 2000). For instance, certain tones have a more central musical function and are more often placed at the beginning and end of any piece of music (Krumhansl, 2000). Consequently, strong expectations are formed, even preattentively, for specific tones in specific positions within a musical piece. These expectations are indexed by the early right anterior negativity (ERAN), an ERP component peaking at about 150 ms after sound onset to out-of-key chords placed at the end of 5-chord cadences, in an experimental condition where subjects were intent in reading a book (Koelsch et al., 2002). In contrast, musical scales create expectations about what categories of events are likely to occur, but not when or in which order. So far no study has uncovered the existence of early ERP effects of musical scale violations. We predicted that pitch deviations from the relational aspects of equal-tempered musical scales should also elicit early negativities outside the focus of attention.

To this aim, we chose a paradigm mimicking a realistic listening condition with pitch deviations from the musical scales, inserted in various locations within unfamiliar unaccompanied melodies and with different levels of attentional load. Specifically, we measured subjects' brain responses in two experiments, one in which they were watching a movie and ignoring the melodies (passive experiment), and another in which they were rating the congruousness of the melodies (active experiment). The melodies included two different kinds of pitch deviances (see Fig. 1). The **out-of-tune** deviance consisted of a tone that was a half semitone interval from the preceding tone, introducing an incongruity from the chromatic scale or tuning of the melody. The **out-of-key** deviance consisted of a tone that was a semitone interval from the preceding tone, placing this tone outside the key of the melody. The **congruous** pitches that served as control comparison were located at corresponding locations in the melodies and were instead at a whole tone or a larger interval from the preceding tone. The other pitches of the melodies all belonged to the respective diatonic scale, thus including several intervals, from the semitone to the octave.

We reasoned that the presence of an attention-independent difference in the brain responses between the incongruous and congruous pitches would support the hypothesis that the brain is able to distinguish tones belonging to scales

Fig. 1 – Three examples of the melodies varied in the three kinds of embedded pitch conditions. The arrows indicate the location of the pitch condition in each melody.

from those that do not. Because the congruous pitches were larger in distance from the previous pitch than the out-of-key incongruities (a semitone) and the out-of-tune ones (a quartertone), larger change-related brain responses to pitch incongruities would support the prediction that pitch relations are automatically encoded according to the musical scale. Otherwise, that is, if the change-related brain responses are more sensitive to pitch distance than scale violation, the reverse should be observed (for a review, see Näätänen et al., 2003): a larger response should be observed for the congruous pitches than for the incongruous ones. By trading the belongingness to the musical scale against pitch distance, we should be able to test the hypothesis that the human brain is sensitive to the stimulus-invariant knowledge of musical scales, and not just to the physical properties of the stimuli. Furthermore, by adopting the two types of deviations, we could test whether the brain would respond differently according to the hierarchy of musical properties violated: the out-of-tune pitch violated the belongingness to the chromatic scale, whereas the out-of-key pitch violated the belongingness to the diatonic scale, a subset of the chromatic scale itself. All subjects were musically untrained

participants, allowing us to probe the neuroarchitecture of implicit musical knowledge.

2. Results

2.1. ERP effects

As Fig. 2 illustrates, in both the passive and active experiments, the congruous pitch elicited a fronto-centrally distributed sharp negative deflection, the N1, peaking on average at 120 ms. Its amplitude did not differ between the experiments [main effect of Experiment: $F(1,8)=0.1$, $p=0.7$]. The ERPs elicited by the congruous pitch instead differed at 380–780 ms between the experiments [main effects of Experiment at 380–480 ms: $F(1,8)=6.4$; $p<0.05$, at 480–580 ms: $F(1,8)=5.5$, $p<0.05$, at 580–680 ms: $F(1,8)=5.1$, $p=0.05$, and at 680–780 ms: $F(1,8)=7.1$, $p<0.05$; interactions Experiment × Frontality at 380–480 ms: $F(4,32)=4.8$, $p<0.05$, and at 480–380 ms: $F(4,32)=2.9$, $p<0.05$; and, finally, interactions Experiment × Frontality × Laterality at 380–480 ms: $F(8,64)=2.5$, $p<0.05$, at 480–580 ms Experiment × Frontality × Laterality: $F(8,64)=2.1$, $p<0.05$, and at
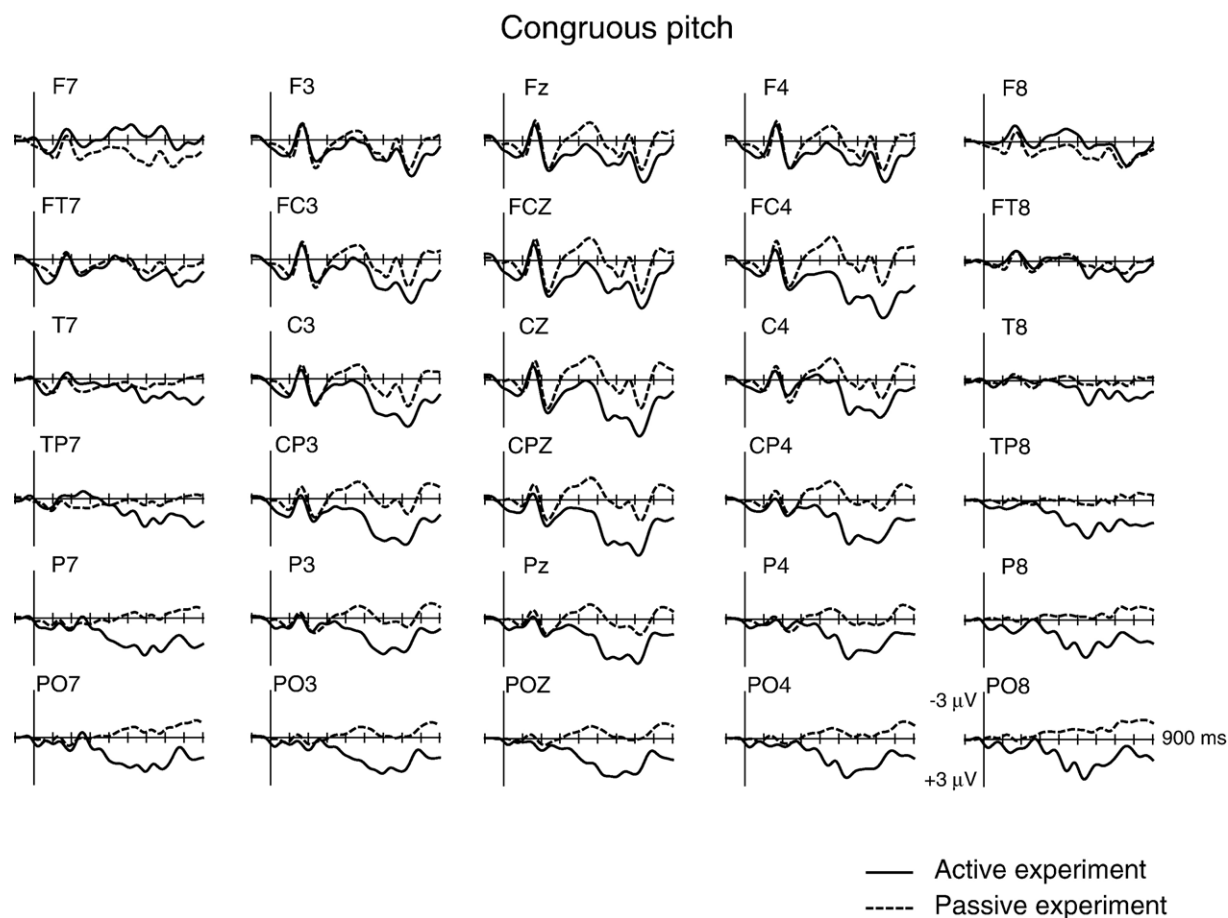
Congruous pitch

Fig. 2 – Grand-average ERPs to the congruous pitch in the passive and active experiments.

580–680 ms: $F(8,64)=2.1$, $p<0.05$]. This effect resulted from the larger long-lasting positive deflection in response to the congruous pitch when presented under the condition of focused attention, i.e., during the active experiment, than when presented during the passive experiment.

As shown by Figs. 3 and 4, in both experiments at the latencies following the N1, a frontally maximal negativity was more pronounced to the out-of-tune pitch, and to a lesser extent also to the out-of-key pitch, as compared with that elicited by the congruous one. This negativity persisted up to about 450 ms in the passive experiment only. For both experiments at later latencies, other peaks were also visible and partially overlapped the ERP responses to the pitch incongruities. Those deflections corresponded to the N1 and P200 elicited by the next tone of the melodies, intervening at 500 ms from the onset of the stimulus of interest.

At 180–280 ms in both the passive and active experiments, the ERPs to the pitch categories differed from each other [main effect of Pitch: $F(2,16)=5.4$, $p<0.02$; without a significant effect of Experiment: $F=1.95$, $p=0.2$, or significant interactions]. In particular, at frontal, fronto-central, and central electrodes, the negativity to the out-of-tune pitch was larger in amplitude than to the other stimuli and the negativity to the out-of-key pitch was larger in amplitude than that to the congruous pitch [interaction Pitch×Frontality: $F(8,64)=11.6$, $p<0.0001$; post hoc

tests: $p<0.01$–0.0001]. No differences were found between the pitches at the parietal and parieto-occipital regions.

At longer latencies, the ERPs to the three pitch conditions differed between the experiments, being more positive – or less negative – in the active than in the passive experiment [main effects of Experiment at 280–380 ms: $F(1,8)=16.6$, $p<0.01$; at 380–480 ms: $F(1,8)=15.1$, $p<0.01$; at 480–580 ms: $F(1,8)=20.9$, $p<0.01$; at 580–680 ms: $F(1,8)=14.1$, $p<0.01$; and at 680–780 ms: $F(1,8)=9.9$, $p<0.05$]. Additionally, at 480–580 ms, an interaction Experiment×Pitch was observed [$F(2,16)=3.9$, $p<0.05$]. Consequently, the analyses were carried out separately for each experiment.

At 280–380 ms, the ERPs to the pitch conditions under the passive experiment differed from each other [main effect of Pitch: $F(2,16)=10$, $p<0.01$, $\varepsilon=0.9$]. In particular, the out-of-tune pitch elicited a larger negativity than the out-of-key and congruous pitches at the frontal, fronto-central, and central regions [interaction Pitch×Frontality: $F(8,64)=8.5$, $p<0.0001$; post hoc tests: $p<0.001$–0.0001], and the out-of-key pitch elicited a larger negativity than the congruous pitch at frontal and fronto-central regions (post hoc tests: $p<0.001$ and 0.05, respectively).

At 380–580 ms in the passive experiment, the negativities to the three pitches differed from each other only at specific electrode locations [interactions Pitch×Frontality at 380–480 ms: $F(8,64)=8.6$, $p<0.0001$; and at 480–580 ms $F(8,64)=$

Passive experiment
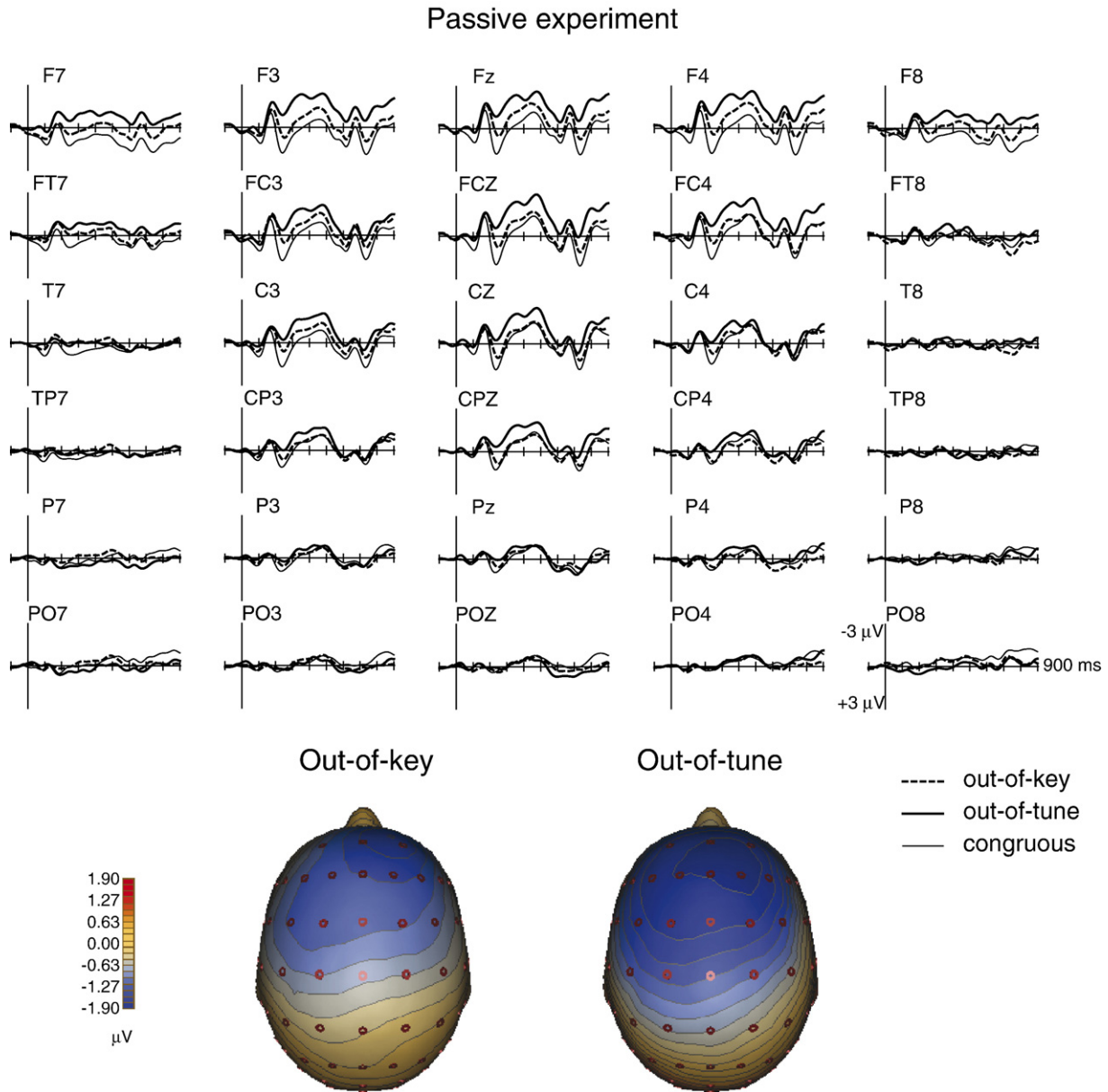
Fig. 3 – Grand-average ERPs to the congruous pitch, out-of-key pitch, and out-of-tune pitch in the passive experiment. The voltage maps are calculated at the early negative frontal peaks of the difference waves (out-of-key minus congruous: 181 ms, and out-of-tune minus congruous: 185 ms).

13.2, $p < 0.0001$]: the negativity to the out-of-tune pitch was even at this long latency larger in amplitude than that to the out-of-key and congruous pitches at frontal, fronto-central, and central regions (post hoc tests: $p < 0.01$–0.0001), and the negativity to the out-of-key pitch was larger than that to the congruous pitch at frontal electrodes only (post hoc tests: $p < 0.1$–0.001). Moreover, at 380–480 ms, the left- and right-hemisphere responses to the three pitches also differed from each other [interaction Pitch×Frontality×Laterality: $F(16,128) = 1.9$, $p < 0.05$]. Separate analyses, including the left and right electrodes of each region of interest, revealed that the negativities were larger in amplitude at the right than at the left hemisphere at the frontal region [main effect of Laterality: $F(1,8) = 6.5$, $p < 0.05$].

At 580–780 ms in the passive experiment, the negativity to the out-of-tune pitch remained larger in amplitude than to the out-of-key and congruous pitch conditions at the frontal and fronto-central regions [interactions Pitch×Frontality at 580–680 ms: $F(8,64) = 6.3$, $p < 0.0001$; post hoc tests: $p < 0.05$–0.0001; and at 680–780 ms: $F(8,64) = 9.3$; $p < 0.0001$; post hoc tests: $p < 0.05$–0.0001], whereas the negativity to the out-of-key pitch was larger in amplitude than to the congruous pitch at the frontal electrodes only (post hoc tests: $p < 0.01$). At 680–780 ms, the incongruous pitches also elicited a more positive potential than did the congruous pitch at the parieto-occipital electrodes [post hoc tests: $p < 0.05$ for both; at this region of interest, the positivity was larger in amplitude over the right than the left hemisphere, as shown by the interaction
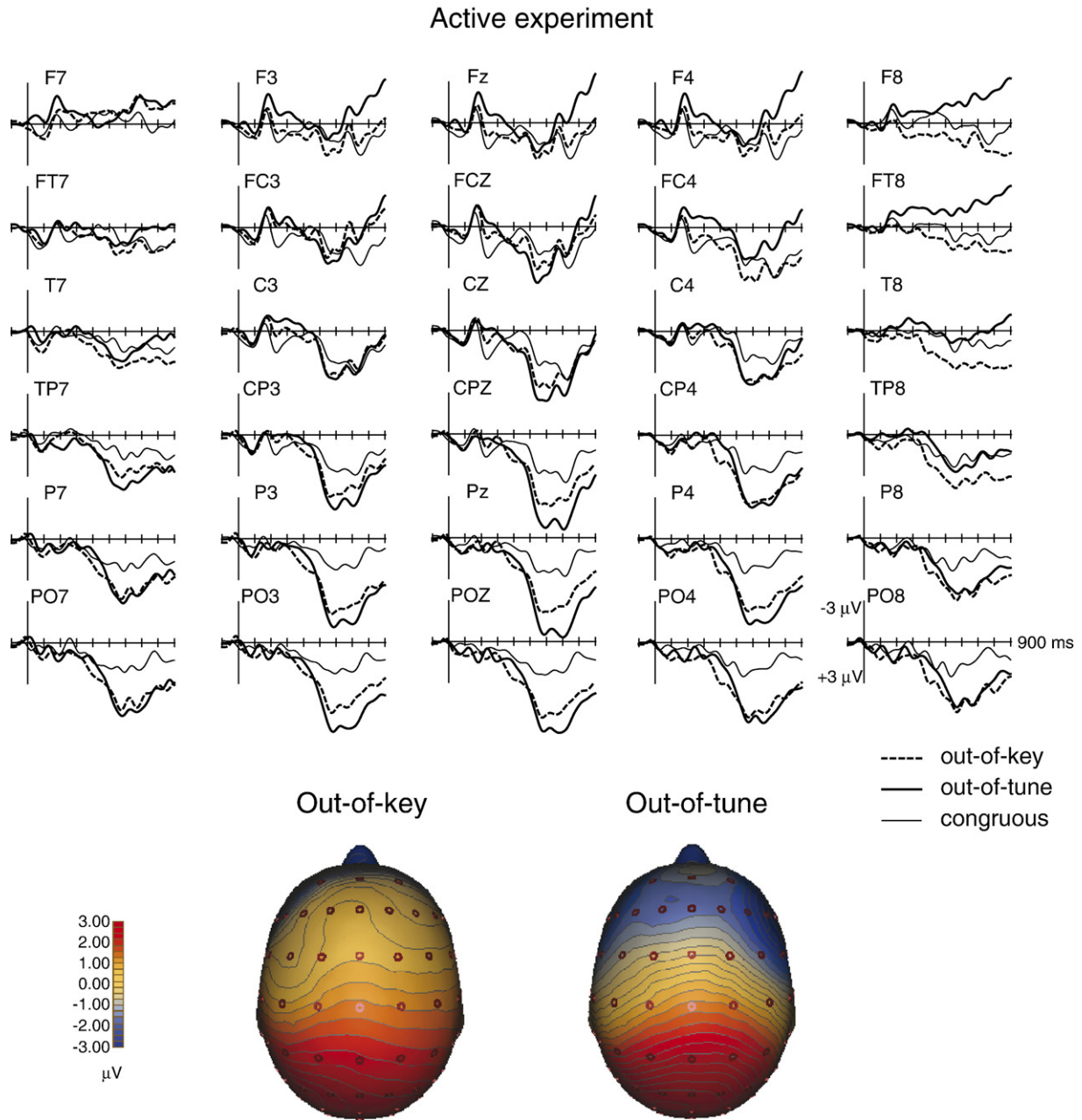
Active experiment

Fig. 4 – Grand-average ERPs to the congruous pitch, out-of-key pitch, and out-of-tune pitch in the active experiment. The voltage maps are calculated at the late positive parieto-occipital peaks of the difference waves (out-of-key minus congruous: 607 ms, and out-of-tune minus congruous: 599 ms).

Pitch×Frontality×Laterality, $F(16,128)=2$, $p<0.05$, and by the main effect of Laterality in the ANOVA carried out on the parieto-occipital amplitudes only, $F(1,8)=7.3$, $p<0.05$].

Turning now to the active experiment, at late latencies following the early negativity to the out-of-tune pitch (which, as reported above, did not differ between experiments), enhanced positive deflections over the parietal and occipital scalp regions (not visible in the passive condition) were elicited by the incongruous out-of-key and out-of-tune pitches as compared with the responses to the congruous pitch (see Fig. 4). These responses resemble the P600 reported in the literature (cf. Besson and Schön, 2003). Detailed analyses

showed that, at 280–480 ms, there were no differences in the neural responses between the three pitches. At the P600 latency range of 480–580 ms, the positivities associated to the out-of-tune and out-of-key pitches were larger than those observed for the congruous pitch [main effect of Pitch: $F(2,16)=4.2$, $p<0.05$, $\varepsilon=0.8$]. This was apparent at all electrodes except for the frontal region [interaction Pitch×Frontality: $F(8,64)=4.1$, $p<0.001$]. At 480–580 ms, we also obtained no difference between the positivities to the out-of-tune and out-of-key pitches. However, the ERP scalp distribution of these responses reveals an earlier differentiation. At 380–480 ms, there was a larger negativity at the right than at the left

hemisphere for the out-of-tune pitch only [interaction Pitch × Frontality × Laterality: $F(16,128)=2.1$, $p<0.05$].

At 580–780 ms, the positivities associated with the out-of-tune and out-of-key pitches were still larger than for the congruous pitch at the parietal and parieto-occipital regions (and also at the central regions for the out-of-tune pitch) [interactions Pitch × Frontality at 580–680 ms: $F(8,64)=12.4$, $p<0.0001$, post hoc tests: $p<0.05$–0.0001; and at 680–780 ms: $F(8,64)=27.6$, $p<0.0001$, post hoc tests: $p<0.05$–0.0001]. Moreover, the out-of-tune pitch elicited a larger positivity than did the out-of-key pitch at the parietal and parieto-occipital regions (post hoc tests: $p<0.05$–0.0001). At the frontal regions, however, the out-of-tune pitch elicited a larger negativity than either the out-of-key deviance or the congruous pitch (post hoc tests: $p<0.05$–0.0001).

2.2.   Source analysis

As shown in Fig. 5, the MCE indicated that the evoked current activity to the out-of-tune pitch in the passive experiment maximal at 185 ms was mainly localized bilaterally in the temporal lobe, but with a larger contribution of the right hemisphere. The local maxima were found in the superior temporal gyrus (Talairach coordinates: $x=-62$, $y=-2$, $z=5$). In the right hemisphere, an additional less strong source occurred in the inferior frontal gyrus (Talairach coordinates: $x=52$, $y=28$, $z=4$).

The MCE calculated for the out-of-key pitch showed that the early negativity maximal at 181 ms was mostly generated

Out-of-key

Out-of-tune

0% █████ 100%
max. %
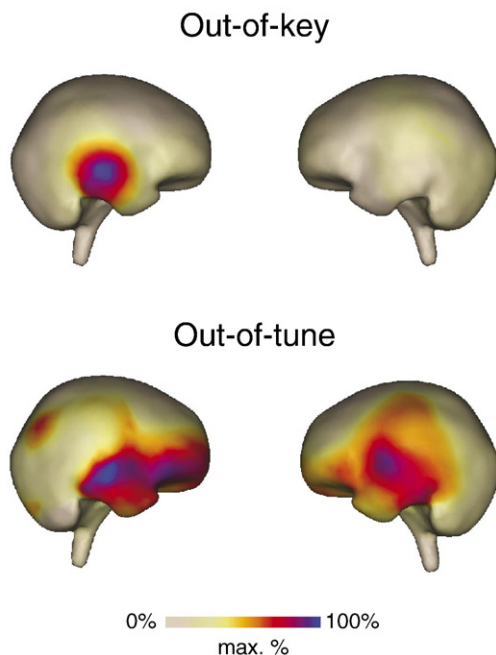
Fig. 5 – Minimum norm current estimation (MCE) images for the early negativities to the out-of-tune and out-of-key pitches in the passive experiment calculated from the grand-averaged referenced-free difference waveforms at the negative frontal peaks within the time window of 180–380 ms (out-of-key minus congruous: 181 ms, and out-of-tune minus congruous: 185 ms). The color code illustrates the strength of the estimated cortical sources in percent calculated for the latency of interest.

Congruousness ratings
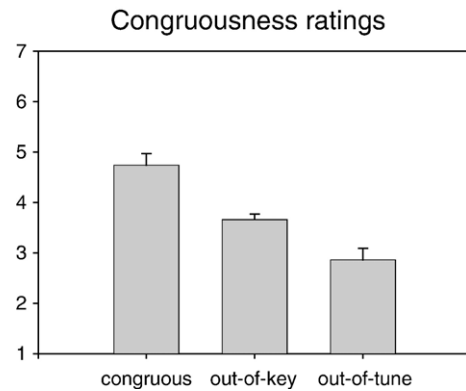
Fig. 6 – Results of subjects' ratings on a 7-point scale obtained during the active experiment. The bars show the standard errors of the mean.

in the right temporal lobe, and particularly in the middle temporal gyrus (Talairach coordinates: $x=67$, $y=-26$, $z=-2$). This result supports our hypothesis that the early negativity to pitch incongruities originates mainly in the secondary auditory cortex (Hall et al., 2003).

2.3.   Subjects' ratings

The ratings of melodic congruousness (Fig. 6) depended on the pitch condition [main effect of Pitch: $F(2,16)=18.5$, $p<0.01$, $\varepsilon=0.6$]. Subjects rated the melodies containing the congruous pitch manipulation as the most congruous (post hoc test: $p<0.01$–0.0001), and the melodies containing the out-of-tune pitch as the most incongruous (post hoc test: $p<0.05$–0.0001). Thus, melodies with the out-of-tune pitch were considered more incongruous than were melodies containing the out-of-key pitch (post hoc test: $p<0.05$).

3.   Discussion

The present study showed that musical scale information is processed automatically by the human brain. More specifically, an early frontal negative neural response was elicited in nonmusicians to musical scale incongruities within a single-voice melody under both the passive and active experiments. Moreover, both the out-of-tune and out-of-key incongruities elicited a negative response. This brain response was larger in amplitude to the out-of-tune than that to the out-of-key pitch, possibly reflecting the larger salience of the first incongruity as compared to the latter.

As reviewed in the Introduction section, pitch expectations can be violated at several levels of the pitch hierarchies in a musical context. In the present paradigm, both the out-of-tune and out-of-key incongruities elicited a negative response, but the brain response to the mistuning was larger in amplitude and more widespread in topography than was the response to the out-of-key pitch. This suggests that despite creating an interval with the preceding tone larger in size (a semitone) than that introduced by the out-of-tune pitch (a quartertone) the out-of-key deviances are less salient than the

out-of-tune deviances. The difference in salience between the two pitch incongruities is further testified by the behavioral responses of the subjects, who rated the melodies with the out-of-tune pitch as more incongruous than the melodies containing the out-of-key pitch. Previous results also demonstrated a larger MMN to the more salient high melodic line as compared to the low one (Fujioka et al., 2005). Furthermore, in our study, the out-of-key pitch deviated from the rule of tonality belongingness, whereas the out-of-tune incongruity deviated from the more general rule of belongingness to the chromatic scale. This latter rule regards all the 12 pitches of Western tonal music, from which the tonalities are formed. The larger negative response to the out-of-tune than to the out-of-key pitch deviation hence confirms and generalizes previous findings by showing the dependence of the automatic auditory cortex functions on the level of salience and the processing demands of the musical scale properties.

An acoustical account of these early negativities is unlikely. It should be noticed that the out-of-tune pitch employed in the present study did not contain any roughness in itself (i.e., amplitude modulations within the sound) nor did it produce any sensory dissonance with other simultaneous sounds since it was played with no harmonic accompaniment. On the other hand, other processes, such as the integration of sounds in sensory memory and effects of interval familiarity, may induce a sensation of dissonance or unpleasantness even with melodic intervals (i.e., not played simultaneously; see, for instance, Moore, 1989; Schellenberg and Trehub, 1996). This dissonance sensation may have been stronger when associated with the quarter tone interval introduced by the out-of-tune pitch than with the semitone interval generated by the out-of-key pitch. However, dissonance seems to produce mainly late positivities in the ERPs (Regnault et al., 2001). In the current study, when presented within the melodic context, the out-of-tune pitches elicited an early negativity during the passive experiment, and the melodies containing those pitches were rated as the most incongruous in the judgment task of the active experiment. In all likelihood, these effects reflect the implicit knowledge of the basic rule of the equal-tempered scale, with the smallest allowed interval being the semitone. This principle was violated by the out-of-tune change that introduced a quartertone (i.e., half semitone) interval within the melody stimulus. Noteworthily, such implicit knowledge seems to be available in musically untrained subjects.

Alternatively, the early negativity to the incongruous pitches might be an example of the ability of the central auditory system to extract and apply complex rules (Tervaniemi et al., 1994; Paavilainen et al., 1999; Wolff and Schröger, 2001; Horvath and Winkler, 2004). For instance, Wolff and Schröger (2001) showed that an infrequent tone repetition elicits an MMN when occurring in a series of tones varying in frequency. Adapted to the present study, the system may apply the rule that adjacent tones are separated by at least one semitone or a whole tone; the deviant introduces instead a smaller frequency change, thus generating an MMN-like brain response. However, this interpretation of the data could only explain the brain reaction to the out-of-tune pitch. The neural response to the out-of-key pitch, instead, cannot be accounted for by a primitive intelligence of the central auditory system

for simple rule extraction (cf. Näätänen et al., 2001) since semitone intervals (e.g., between the seventh and eight tones of the diatonic scale) occurred in several of the melodies. In other words, the auditory system could not simply compute the presence of a deviant by extracting the rule that melodies contained only pitch distances equal to or larger than a whole tone; it rather needed to compare the incoming sounds with the long-term neural traces for musical pitch relations stored as neural assemblies in the cortex.

Memory representations for repeated or meaningful stimuli of the environment are supposed to be stored in the regions of the brain where their initial processing also takes place, i.e., in the sensory cortices (Weinberger, 2004; Destexhe and Marder, 2004). A comparison process which occurs automatically in the brain compares the incoming sounds with the memory traces present in the auditory cortex, as indexed by the MMN component of the ERPs, occurring as early as 150–250 ms from the onset of the sound discrepant with the stored neural traces. In the traditional MMN paradigms, the neural trace is of a sensory nature; that is, it is formed during the experimental session by repeating specific sound parameters or simple invariances of the sound stimulation (such as the pitch direction in tone pairs; Saarinen et al., 1992). When the repeated sounds or sound relations are familiar, the MMN is enhanced, indexing the automatic activation of long-term memory traces for those sounds or sound relations in the auditory cortex (Näätänen et al., 1997; Pulvermüller et al., 2001; Schröger et al., 2004). In the present study, where the sound repetitions were minimized, the brain response to the pitch violation was solely the result of the comparison of the incoming pitch with the long-term traces for the musical scale properties rather than with the sensory memory traces for the invariant pitches presented during the experimental session. Specifically, the incongruous pitches did not match the permanent neural traces for the pitch relations of the musical scale in the human brain activated by the preceding melody context. In other words, at this early stage, the comparison process did not use individual sounds but the scale structure as its reference point. As an end product, we could observe the present early negativity, closely corresponding to the MMN, in response to the out-of-tune and out-of-key pitches.

The source analysis localized the present MMN to pitch incongruities mainly in the supratemporal lobe, corresponding to the secondary auditory cortex, with a predominant contribution from the right hemisphere (a weaker source was also observed in the frontal cortex). This finding is in line with previous brain imaging and neuropsychological evidence associating the secondary auditory cortex (in particular the right-hemispheric one) with the processing of the contour properties of unfamiliar melodies, as contrasted with the primary auditory cortex analyzing features of isolated sounds only (Milner, 1962; Samson and Zatorre, 1988; Johnsrude et al., 2000; Patterson et al., 2002). Our data thus suggest that the melodies were automatically modeled by the secondary auditory cortex as based on the pitch relations of the musical scale, much like linguistic stimuli are automatically categorized according to their phonological content (see below). On the basis of our findings, we propose that the efficient

computation of the pitch relations of the diatonic musical scale is based on the long-term practice of the neural networks of the auditory cortex to the rules of Western tonal music in listeners acculturated with it.

The present results thus mirror earlier findings from the linguistic domain, where the brain is preattentively sensitive to abstract phonological information (Phillips et al., 2000; Näätänen, 2001; Kujala et al., 2002). Like phonemes within a word context, a given pitch is perceived as out-of-key or out-of-tune only within the melodic context of the adjacent tones. Even if it seems to occur as early as at the preattentive level, the detection of such contextual information is not a simple feat but requires abstract relational knowledge and the use of long-term memory processes for the computation of the relevant comparisons (cf. Näätänen et al., 2001).

Previously, preattentive processing of harmonic relations in the brain has been studied with chord cadences (Koelsch et al., 2002). The results showed that an ERAN (Koelsch et al., 2002) was elicited by the out-of-key Neapolitan chord ending chord cadences even when subjects were intent in reading a book while ignoring the sounds. Our results confirm and extend these findings to the domain of the musical scale relations in melodies. On the other hand, harmony processing, indexed by the ERAN, is not fully automatic since the ERAN amplitude decreases when sounds are outside attentional focus (Loui et al., 2005). In the current study, instead, the early negativity to the pitch incongruities was not modulated by subjects' attention, thus suggesting that musical scale processing is fully independent of attentional resources. Moreover, the present negativity was generated mainly in the superior temporal lobe (with a predominant contribution of the right hemisphere and with a secondary possible source in the frontal cortex), as showed by the scalp distribution of the ERPs, maximal at frontal and fronto-central electrodes and reversing their potential at temporo-mastoidal sites, and by the electrical source analysis. In contrast, the electric ERAN peaks at frontal and fronto-temporal scalp regions (see, e.g., Koelsch et al., 2000; Koelsch and Mulder, 2002) and its magnetic counterpart is generated mainly in the right and left inferior frontal cortices, as evidenced by dipole modeling (Maess et al., 2001). In light of previous studies concerning the role of these brain regions (e.g., Smith and Jonides, 1998; Grodzinski, 2000; Näätänen et al., 2001; Korzyukov et al., 2003), we propose that the different localization of ERAN and the present MMN may result from the different types of violations investigated: in the ERAN studies, the Neapolitan chord, especially when placed at the end of cadences, violates the rules of harmony concerning the order of sound events within a structure (cf. Snyder, 2000), whereas in our study the out-of-tune and the out-of-key pitches inserted in various locations within the melodies violate the rules of belongingness to the musical scale. In other words, the hierarchical rules of harmony that require the combination of several musical units into meaningful complex representations ordered in time tend to be associated with the frontal regions of the brain (Maess et al., 2001), whereas the non-hierarchical relational properties of the musical scale seem to be mostly extracted in auditory cortex areas (cf. Näätänen et al., 2001). Consequently, on the basis of our source analysis, we propose that musical scale

processing is more analogous to phoneme extraction in the domain of language, and thus with the MMN concept, than with syntax, in contrast with what has been argued for the ERAN brain response (Koelsch et al., 2000; Koelsch and Siebel, 2005).

Another discrepancy in relation to the Koelsch et al.'s (2002) study lies in the relatively small amplitude of the present early negative response to the out-of-key pitch, probably due to the different paradigms used. In order to investigate preattentive harmony processing, Koelsch et al. (2002) opted for a relatively repetitive musical context in which chords, while transposed over different keys and played in various registers, were isochronously presented (with sounds occurring equi-distantly in time). In the present paradigm, the musical context varied in rhythm, the melodies were played over different keys, and the moment at which a pitch incongruity occurred varied. Consequently, the occurrence of an incongruity could not be easily predicted, which may have increased uncertainty regarding expectations (Näätänen, 1970), and thus decreased the power of violation-elicited neural activations.

In the active experiment, we additionally observed the attention-related P600 component, which was larger in amplitude to both the out-of-tune and the out-of-key pitches as compared with that elicited by the congruous pitches, not differing from each other in the 480–580 ms time window. This lack of amplitude difference in the initial P600 between the more salient out-of-tune pitch and the less salient key violation indicates that additional neural resources were recruited in order to attentively identify and integrate the out-of-key pitch in the melody context. Thus, we suggest that melody processing is completed with the aid of focused attention, which presumably contributes to fully integrate musical scale information into the ongoing pitch analysis. The P600 findings in response to out-of-key and out-of-tune pitches within melodies support and generalize previous findings to an ongoing listening situation showing that pitch incongruities placed at the end of melodies elicit a parietally maximal P600 (Besson et al., 1994; Besson and Faita, 1995; cf. also Besson and Schön, 2003). Contrary to the majority of the previous experiments on music processing (e.g., Paller et al., 1992; Besson et al., 1994; Besson and Faita, 1995; Janata, 1995; Hantz et al., 1997; Patel et al., 1998; Koelsch et al., 2000; Regnault et al., 2001; Tervaniemi et al., 2003), in the present study, subjects could hardly predict the moment in time they should expect an incongruous event. This made the experimental situation more similar to a common music listening situation. Nevertheless, the congruousness judgments given to the wholeness of the melody showed that a single pitch in a variable place inside the melody is sufficient to affect the subjects' ratings in a significant way.

In conclusion, we propose that the human brain possesses mechanisms to extract relational aspects of the sounds of the musical scale without a need for focused attention, but later calling into play attentional resources for fully integrated, conscious access. Consequently, melody processing seems to be driven by fast automatic processes occurring in the secondary auditory cortex. Such processes are based on the knowledge available in the brains of the majority of listeners, i.e., in subjects without any musical

education, who have implicitly learned musical properties through everyday passive exposure to music.

4. Experimental procedures

4.1. Subjects

Nine healthy right-handed students (mean age 23±3, 4 females) with no formal musical education were tested with electroencephalography (EEG). They gave formal written consent to participate in the study. The experimental procedures followed the guidelines reported in the Declaration of Helsinki.

4.2. Materials

In both the active and passive experiments, subjects were presented with 40 unfamiliar melodies of approximately 6 s in duration (see Fig. 1). The melodies, composed for experimental purposes at the University of Montreal, differed from one another in pitch and rhythmic content. They were played in different keys and were structured according to the rules of the Western tonal system. Half of the melodies were written in binary tempi and half in ternary tempi. The metronome was set within a range of 30 to 240 beats per minute, with the majority of the melodies played at 120 beats per minute. In half of the melodies, a pitch manipulation occurred on the strong beat of the third bar and lasted about 500 ms. Since the melodies consisted of a different number of tones and different rhythms, the manipulation was introduced in varying locations within 2–4 s after melody onset. The pitch manipulations were of two kinds, thus each occurring in 25% of the melodies: an out-of-key pitch (by a semitone interval from the preceding pitch) introduced a deviation from the key of the melody, and an out-of-tune pitch (by a quartertone interval from the preceding in-key pitch) introduced a deviation from the chromatic scale or the tuning of the melody. Those incongruous pitches were compared with the congruous in-key pitches occurring at corresponding locations in the melodies. In the current paradigm, the out-of-key and out-of-tune pitches were also incongruous as compared to the other pitches of the diatonic melodies. The actual probability of occurrence of the pitch deviances was hence much less than 25%. It is worth pointing out that the out-of-tune pitch had a very similar frequency spectrum to the other pitches (when heard in isolation, it sounded perfectly consonant), but its distance from the preceding pitch was half a semitone (e.g., a pitch half way between C and C#), thus producing a small interval not commonly used in Western tonal music. In order to warrant precise time locking of the ERP, the onset of the critical pitch was marked by way of a careful inspection of the auditory and spectral signal.

Each melody was presented 4 times: twice with different congruous pitches and twice with different incongruous pitches. The melodies were computer-generated and played on three different instruments, a nylon string guitar, a clarinet, or a jazz guitar (on a Roland SC 50 sound canvas). In total, in the study, 480 melodies were presented. In summary, the contour, rhythm, pitch level, and timbre of each melody

varied, thus minimizing their surface-level (or pitch-level) invariance. Instead, the pitch invariance was related to the belongingness to the equal-tempered musical scale and to the specific key in which the melody was composed.

4.3. Procedure

The EEG measurements were performed at the University of Montreal in a single session lasting about 4 h. The melodies were binaurally presented through Sennheiser HD450 headphones in a quiet room, at an intensity level of 70 dB SPL, and with an interstimulus interval (ISI) of 4 s for both the experiments.

In the passive experiment, subjects were presented with 480 melodies and asked to watch a soundless DVD movie with subtitles while ignoring the sounds. After the movie, subjects were given a break of 30 min during which they had refreshment and were allowed to move. In the active experiment, always administered after the passive experiment, the melodies were played with one of the previously used timbres (160 trials) while subjects were performing a paper and pencil test. In this behavioral test, subjects judged the congruousness and incongruousness of each melody. They were requested to judge whether the melodies contained a 'wrong' pitch on a 7-point scale, in which 1 meant very incongruous, 4 neutral, and 7 very congruous. Importantly, subjects were not informed of the location in the melody in which the pitch manipulation would occur. Subjects received 4 practice trials without feedback before performing the task. The results obtained in the behavioral test were analyzed with a 1-way repeated-measure ANOVA (pitch: congruous, out-of-key, out-of-tune).

4.4. EEG recordings

The EEG was recorded with an InstEP amplifier from 62 tin electrodes (Electrocap International, Inc.®) arranged on the scalp according to the extended 10–20 international system appended by intermediate positions and by the left and right mastoids. All electrodes were referenced to the electrode placed on the nose. Horizontal and vertical electrooculograms (EOG) were bipolarly monitored with electrodes placed above and below the right eye and at the left and right eye canthi. The EEG and EOG were amplified (bio-electric amplifier by SA Instrumentation; 256 Hz sampling rate) with a bandpass of 0.15 to 50 Hz.

4.5. Data analysis

Continuous EEG records were divided into epochs starting 100 ms before and ending 900 ms after the onset of the manipulated pitch. EEG epochs contaminated by blinks or eye-movement artifacts were corrected by a dynamic regression procedure of the EOG on the EEG in the frequency domain (Woestenburg et al., 1983). Epochs with a signal change exceeding ±150 μV at any EEG electrode were rejected from averaging. ERP averaging was performed without regard to the subject's behavioral response. ERPs were offline filtered digitally (bandpass 0.5–25 Hz at 24 dB/octave), re-referenced to the algebraic mean of both mastoids in order to improve the

signal to noise ratio, and quantified for each subject in each pitch condition and for each electrode (Neuroscan Ltd., El Paso, Tx., Edit 4.2).

In both passive and active experiments, in order to test any differences in the neural responses to the two pitch incongruities as compared with those to the congruous pitch, we quantified the amplitudes and latencies of the ERPs to the three stimulus categories from 15 electrodes (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, P3, Pz, P4, PO3, POz, PO4) at sliding latency windows of 100 ms, starting from 180 ms. The latency of 180 ms was chosen as starting point for the statistical testing because it approximately corresponds to the latency in which the ERPs to temporally and spectrally complex deviant and standard stimuli start to diverge, as described in the literature (Näätänen et al., 1993; Paavilainen et al., 1999; Tervaniemi et al., 2001; Brattico et al., 2002; van Zuijen et al., 2004), and because the visual inspection of the grand-average difference waveforms (in which the ERPs to the incongruous pitches were subtracted from the ERPs to the congruous pitches) revealed the maximal effects shortly after this latency. The procedure of doing the statistics for subsequent latency windows was adopted since the negativity to pitch incongruities in the passive experiment and the positivity to pitch incongruities in the active experiment were long lasting. This procedure is also consistent with the literature: ERP components associated to complex auditory or cognitive processes are most commonly analyzed over wide time windows (see, for instance, Näätänen et al., 1993; Hahne and Friederici, 1999; Tervaniemi et al., 2001; Schön and Besson, 2005; Nicholson et al., 2006; Nan et al., 2006). Moreover, we wished to test whether the long-lasting ERP deflections observed in our study varied in scalp distribution and hence in functional significance at different latency ranges.

As visible from the grand-average waveforms (Figs. 2–4), while the late negativities for the passive experiment and the late positivities for the active experiment had long latencies, the first negativity for all pitch categories, corresponding to the N1 component of the ERPs, had a sharp peak. Consequently, only for the N1, we measured the mean amplitudes from the 40-ms window around the peaks identified from the grand-average waveforms.

The mean amplitudes of the ERP components of interest were then compared with repeated-measure ANOVAs including, when appropriate, Experiment (passive, active), Pitch (congruous, out-of-key, out-of-tune), Frontality (F-line, FC-line, C-line, P-line, PO-line), and Laterality (Left, Middle, Right) as factors. In all statistical analyses, type I errors were controlled for by decreasing the degrees of freedom with the Greenhouse–Geisser epsilon (the original degrees of freedom for all analyses are reported throughout the paper). *Post hoc* tests were conducted by Fisher's least-significant difference (LSD) comparisons.

4.6.   *Source analysis*

To assess the possible source location of the early negativities obtained under the passive experiment to the pitch incongruities, we calculated L2 minimum-norm current estimation (MCE) (Hämäläinen and Ilmoniemi, 1984; Hämäläinen and Ilmoniemi, 1994) by using the Brain Electrical Source Analysis

software (BESA 5.6.01). MCE calculates a distributed current image at each time sample on the basis of the potential distribution recorded as the smallest amplitude of the overall activity (Hämäläinen and Ilmoniemi, 1984). MCE was preferred here because it has relatively good localization accuracy, and it requires minimal assumptions about the activity distribution, as compared with the dipole method which, for instance, confines the neural activity to point-like sources (Komssi et al., 2004). In our analysis, no a priori knowledge about the source location was introduced, apart from restricting the source to the cortical surface.

Since MCE is very sensitive to the noise level in the signal, we performed it on the grand-average reference-free difference waveforms where responses to the congruous pitch were subtracted from those to the incongruous pitch. The high-pass filter of 0.5 Hz (24 dB/octave) and the low-pass filter of 10 Hz (24 dB/octave) were also applied to increase the signal to noise ratio of the grand-average waveforms (see, e.g., Sinkkonen and Tervaniemi, 2000). The MCE images were then computed as regional sources evenly distributed on 1420 standard locations 10% and 30% below the smoothed standard brain surface of the BESA software. For this computation, we used spatio-temporal weighting according to Dale and Sereno (1993), which assigns more weight to sources that are assumed to contribute more to the data recorded. The MCE images were finally drawn from the difference waveforms at the latency of the negative peak recorded from the frontal electrodes within the 180–380 ms time window.

R E F E R E N C E S

Besson, M., Faita, F., 1995. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. J. Exp. Psychol. Hum. Percept. Perform. 21, 1278–1296.

Besson, M., Schön, D., 2003. Comparison between language and music. In: Peretz, I., Zatorre, R. (Eds.), The Cognitive Neuroscience of Music. Oxford Univ. Press, New York, pp. 269–293.

Besson, M., Faita, F., Requin, J., 1994. Brain waves associated with musical incongruities differ for musicians and non-musicians. Neurosci. Lett. 168, 101–105.

Brattico, E., Näätänen, R., Verma, T., Välimäki, V., Tervaniemi, M., 2000. Processing of musical intervals in the central auditory system: an event-related potential (ERP) study on sensory consonance. Proceedings of the Sixth International Conference on Music Perception and Cognition. Keele University, Keel, pp. 1110–1119.

Brattico, E., Näätänen, R., Tervaniemi, M., 2001. Context effects on

pitch perception in musicians and non-musicians: evidence from ERP recordings. Music Percept. 19, 1–24.

Brattico, E., Winkler, I., Näätänen, R., Paavilainen, P., Tervaniemi, M., 2002. Simultaneous storage of two complex temporal sound patterns in the human auditory sensory memory. NeuroReport 13, 1747–1751.

Bregman, A.S., 1990. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press, Cambridge, MA.

Coull, J.T., 1998. Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology. Prog. Neurobiol. 55, 343–361.

Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with fMRI with cortical surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176.

Destexhe, A., Marder, E., 2004. Plasticity in single neuron and circuit computations. Nature 431, 789–795.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., Pantev, C., 2005. Automatic encoding of polyphonic melodies in musicians and nonmusicians. J. Cogn. Neurosci. 17, 1578–1592.

Grodzinski, Y., 2000. The neurology of syntax: language use without Broca's area. Behav. Brain Sci. 23, 1–71.

Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. J. Cogn. Neurosci. 11, 194–205.

Hall, D.A., Hart, H.C., Johnsrude, I.S., 2003. Relationships between human auditory cortical structure and function. Audiol. Neurootol. 8, 1–18.

Hämäläinen, M.S., Ilmoniemi, R.J., 1984. Interpreting measured magnetic fields of the brain: minimum norm estimates of current distributions, Vol. Technical Report TKK-F-A559, Helsinki University of Technology, Helsinki.

Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42.

Hantz, E.C., Kreilick, K.G., Kananen, W., Swartz, K.P., 1997. Neural responses to melodic and harmonic closure: an event-related-potential study. Music Percept. 15, 69–98.

Horvath, J., Winkler, I., 2004. How the human auditory system treats repetition amongst change. Neurosci. Lett. 368, 157–161.

Janata, P., 1995. ERP measures assay the degree of expectancy violation of harmonic contexts in music. J. Cogn. Neurosci. 7, 153–164.

Johnsrude, I.S., Penhune, V.B., Zatorre, R.J., 2000. Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123, 155–163.

Koelsch, S., Mulder, J., 2002. Electric brain responses to inappropriate harmonies during listening to expressive music. Clin. Neurophysiol. 113, 862–869.

Koelsch, S., Siebel, W.A., 2005. Towards a neural basis of music perception. Trends Cogn. Sci. 9, 578–584.

Koelsch, S., Gunter, T.C., Friederici, A.D., Schröger, E., 2000. Brain indices of music processing: 'non-musicians' are musical. J. Cogn. Neurosci. 12, 520–541.

Koelsch, S., Schröger, E., Gunter, T.C., 2002. Music matters: preattentive musicality of the human brain. Psychophysiology 39, 38–48.

Komssi, S., Huttunen, J., Aronen, H.J., Ilmoniemi, R.J., 2004. EEG minimum-norm estimation compared with MEG dipole fitting in the localization of somatosensory sources at S1. Clin. Neurophysiol. 115, 534–542.

Korzyukov, O.A., Winkler, I., Gumenyuk, V.I., Alho, K., 2003. Processing abstract auditory features in the human auditory cortex. NeuroImage 20, 2245–2258.

Krumhansl, C.L., 2000. Rhythm and pitch in music cognition. Psychol. Bull. 126, 159–179.

Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. 5, 831–843.

Kujala, A., Alho, K., Valle, S., Sivonen, P., Ilmoniemi, R.J., Alku, P., Näätänen, R., 2002. Context modulates processing of speech sounds in the right auditory cortex of human subjects. Neurosci. Lett. 331, 91–94.

Loui, P., Grent-'t-Jong, T., Torpey, D., Woldorff, M., 2005. Effects of attention on the neural processing of harmonic syntax in Western music. Cogn. Brain Res. 25, 678–687.

Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area: an MEG study. Nat. Neurosci. 4, 540–545.

Milner, B., 1962. Laterality effects in audition. In: Mountcastle, V.B. (Ed.), Interhemispheric Relations and Cerebral Dominance. Johns Hopkins, Baltimore, pp. 177–195.

Moore, B.C.J., 1989. An Introduction to the Psychology of Hearing. Academic Press, London.

Müller, B.W., Jüptner, M., Jentzen, W., Müller, P.S., 2002. Cortical activation to auditory mismatch elicited by frequent deviant and complex novel sounds: a PET study. NeuroImage 17, 231–239.

Näätänen, R., 1970. The diminishing time—uncertainty with the lapse of time after the warning signal in reaction-time experiments with varying fore-periods. Acta Psychol. 34, 399–419.

Näätänen, R., 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology 38, 1–21.

Näätänen, R., Winkler, I., 1999. The concept of auditory stimulus representation in cognitive neuroscience. Psychol. Bull. 125, 826–859.

Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42, 313–329.

Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., Paavilainen, P., 1993. Development of a memory trace for a complex sound in the human brain. NeuroReport 4, 503–506.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385, 432–434.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., Winkler, I., 2001. Pre-attentive cognitive processing ('primitive intelligence') in the auditory cortex as revealed by the mismatch negativity (MMN). Trends Neurosci. 24, 283–288.

Näätänen, R., Brattico, E., Tervaniemi, M., 2003. Mismatch negativity (MMN): a probe to auditory cognition and perception in basic and clinical research. In: Zani, A., Mado Proverbio, A. (Eds.), The Cognitive Electrophysiology of Mind and Brain. Academic Press, San Diego, pp. 343–355.

Nan, Y., Knösche, T.R., Friederici, A.D., 2006. The perception of musical phrase structure: a cross-cultural ERP study. Brain Res. 1094, 179–191.

Nicholson, R., Karayanidis, F., Bumak, E., Poboka, D., Michie, P.T., 2006. ERPs dissociate the effects of switching task sets and task cues. Brain Res. 1095, 107–123.

Paavilainen, P., Jaramillo, M., Näätänen, R., Winkler, I., 1999. Neuronal populations in the human brain extracting invariant relationships from acoustic variance. Neurosci. Lett. 265, 179–182.

Paller, K.A., McCarthy, G., Wood, C.C., 1992. Event-related potentials elicited by deviant endings to melodies. Psychophysiology 29, 202–206.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. J. Cogn. Neurosci. 10, 717–733.

Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., Griffiths, T.D., 2002. The processing of temporal pitch and melody information in auditory cortex. Neuron 36, 767–776.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., Roberts, T., 2000. Auditory cortex accesses phonological categories: an MEG mismatch study. J. Cogn. Neurosci. 12, 1038–1055.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Ilmoniemi, R.J., Näätänen, R., 2001. Memory traces for words as revealed by the mismatch negativity. NeuroImage 14, 607–616.

Regnault, P., Bigand, E., Besson, M., 2001. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event-related brain potentials. J. Cogn. Neurosci. 13, 241–255.

Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M., Näätänen, R., 1992. Representation of abstract attributes of auditory stimuli in human brain. NeuroReport 3, 1149–1151.

Samson, S., Zatorre, R.J., 1988. Melodic and harmonic discrimination following unilateral cerebral excision. Brain Cogn. 7, 348–360.

Schellenberg, E.G., Trehub, S.E., 1996. Natural musical intervals: evidence from infant listeners. Psychol. Sci. 7, 272–277.

Schmithorst, V.J., 2005. Separate cortical networks involved in music perception: preliminary functional MRI evidence for modularity of music processing. NeuroImage 25, 444–451.

Schön, D., Besson, M., 2005. Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. J. Cogn. Neurosci. 17, 694–705.

Schröger, E., 1994. An event-related potential study of sensory representations of unfamiliar tonal patterns. Psychophysiology 31, 175–181.

Schröger, E., Tervaniemi, M., Huotilainen, M., 2004. Bottom–up and top–down flows of information within auditory memory: electrophysiological evidence. In: Kaernbach, C., Schröger, E., Müller, H. (Eds.), Psychophysics Beyond Sensation: Laws and Invariants of Human Cognition. Erlbaum, Hillsdale, pp. 389–407.

Sinkkonen, J., Tervaniemi, M., 2000. Towards optimal recording and analysis of the mismatch negativity. Audiol. Neurootol. 5, 235–246.

Smith, E.E., Jonides, J., 1998. Neuroimaging analyses of human working memory. Proc. Natl. Acad. Sci. U. S. A. 95, 12061–12068.

Snyder, B., 2000. Music and Memory. MIT Press, Cambridge, MA.

Tervaniemi, M., Brattico, E., 2004. From sounds to music: towards understanding the neural basis of musical sound perception. J. Conscious. Stud. 11, 9–27.

Tervaniemi, M., Maury, S., Näätänen, R., 1994. Neural representations of abstract features in the human brain as reflected by the mismatch negativity. NeuroReport 5, 844–846.

Tervaniemi, M., Medvedev, S.V., Alho, K., Pakhomov, S.V., Roudas, M.S., van Zuijen, T., Näätänen, R., 2000. Lateralized automatic auditory processing of phonetic versus musical information: a PET study. Hum. Brain Mapp. 10, 74–79.

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R.J., Näätänen, R., 2001. Superior formation of cortical memory traces of melodic patterns in musicians. Learn. Mem. 8, 295–300.

Tervaniemi, M., Huotilainen, M., Brattico, E., Reinikainen, K., Ilmoniemi, R.J., Alho, K., 2003. Event-related potentials to expectancy violation in musical context. Music. Sci. 7, 241–261.

Tillmann, B., Bharucha, J.J., Bigand, E., 2000. Implicit learning of tonality: a self-organizing approach. Psychol. Rev. 107, 885–913.

Trainor, L.J., Trehub, S.E., 1993. Musical context effects in infants and adults: key distance. J. Exp. Psychol. Hum. Percept. Perform. 19, 615–626.

Trainor, L.J., McDonald, K.L., Alain, C., 2002. Automatic and controlled processing of melodic contour and interval information measured by electric brain activity. J. Cogn. Neurosci. 14, 430–442.

Trehub, S.E., Thorpe, L.A., Trainor, L.J., 1990. Infants' perception of 'good' and 'bad' melodies. Psychomusicology 9, 5–19.

van Zuijen, T.L., Sussman, E., Winkler, I., Näätänen, R., Tervaniemi, M., 2004. Grouping of sequential sounds—An event-related potential study comparing musicians and nonmusicians. J. Cogn. Neurosci. 16, 331–338.

Velmans, M., 1991. Is human information processing conscious? Behav. Brain Sci. 14, 651–726.

Weinberger, N.M., 2004. Specific long-term memory traces in primary auditory cortex. Nat. Rev., Neurosci. 5, 279–290.

Woestenburg, J.C., Verbaten, M.N., Slangen, J.L., 1983. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. Biol. Psychol. 16, 127–147.

Wolff, C., Schröger, E., 2001. Activation of the auditory pre-attentive change detection system by tone repetitions with fast stimulation rate. Cogn. Brain Res. 10, 323–327.