

Electrophysiological evidence for a specific neural correlate of musical violation expectation in primary-school children



Clara E. James^{a,b,c,*}, Donato M. Cereghetti^{b,d}, Elodie Roulet Tribes^b, Mathias S. Oechslin^{b,e}

^a HES-SO University of Applied Sciences and Arts Western Switzerland, School of Health Sciences, Geneva, Switzerland

^b Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland

^c Geneva Neuroscience Center, University of Geneva, Geneva, Switzerland

^d Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

^e International Normal Aging and Plasticity Imaging Center (INAPIC), University of Zurich, Zurich, Switzerland

ARTICLE INFO

Article history:

Accepted 20 September 2014

Available online 30 September 2014

Keywords:

Musical expectation violation

ERP

Children

Pre-motor

Working memory

ABSTRACT

The majority of studies on music processing in children used simple musical stimuli. Here, primary schoolchildren judged the appropriateness of musical closure in expressive polyphone music, while high-density electroencephalography was recorded. Stimuli ended either regularly or contained refined in-key harmonic transgressions at closure. The children discriminated the transgressions well above chance. Regular and transgressed endings evoked opposite scalp voltage configurations peaking around 400 ms after stimulus onset with bilateral frontal negativity for regular and centro-posterior negativity (CPN) for transgressed endings. A positive correlation could be established between strength of the CPN response and rater sensitivity (d'). We also investigated whether the capacity to discriminate the transgressions was supported by auditory domain specific or general cognitive mechanisms, and found that working memory capacity predicted transgression discrimination. Latency and distribution of the CPN are reminiscent of the N400, typically observed in response to semantic incongruities in language. Therefore our observation is intriguing, as the CPN occurred here within an intra-musical context, without any symbols referring to the external world. Moreover, the harmonic in-key transgressions that we implemented may be considered syntactical as they transgress structural rules. Such structural incongruities in music are typically followed by an early right anterior negativity (ERAN) and an N5, but not so here. Putative contributive sources of the CPN were localized in left pre-motor, mid-posterior cingulate and superior parietal regions of the brain that can be linked to integration processing. These results suggest that, at least in children, processing of syntax and meaning may coincide in complex intra-musical contexts.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Western tonal music relies on a rule system of hierarchical pitch relationships around a key “center” or tonic. In order to study tonal expectation violation, unexpected pitches that transgress this rule system are embedded in musical contexts. This has been done either in monophonic material, tone sequences or melodies (Besson and Faïta, 1995; Nittono et al., 2000; Trainor and Trehub, 1994), or in harmonic progressions, the latter either in chord sequences (Bigand et al., 2003; Koelsch et al., 2001, 2007; Regnault et al., 2001) or in complex polyphone music (James et al., 2008; Koelsch et al., 2008; Koelsch and Mulder, 2002; Oechslin et al., 2013b; Steinbeis et al., 2006). If these transgressions are in-key or in a closely related key, they are generally considered syntactical, because they transgress mainly structural or grammatical rules (James et al., 2008; Koelsch,

2011; Koelsch et al., 2007; Patel, 2003; Tillmann et al., 2006). Such is the case of the stimuli we used in the present investigation.

Brain substrates processing musical tonal expectation violation have been shown to overlap, neighbor or mirror in the contralateral (right) hemisphere those for syntax processing in language in the left hemisphere (Friederici, 2002), in particular in the inferior frontal cortex (Brown et al., 2006; Maess et al., 2001; Oechslin et al., 2013b; Tillmann et al., 2006). A plausible observation, as both language and music articulate discrete elements in embedded hierarchical structures of increasing complexity (James, 2012; Schon and Francois, 2011). Like in language, not only local but also long-distance structural dependencies are essential for syntactic processing in musical contexts, for instance regarding musical closure at the end of a phrase (James et al., 2008; Koelsch et al., 2013; Steinbeis et al., 2006).

That music also may convey meaning is widely accepted (Koelsch, 2011; Koelsch et al., 2004; Meyer, 1956; Swain, 1997), however, the nature of meaning in intra-musical contexts is intrinsically different from that in language, not concrete and varying between individuals even of the same culture (Lamont and Webb, 2010). Intra-musical

* Corresponding author at: HES-SO University of Applied Sciences and Arts Western Switzerland, School of Health Sciences, Geneva, Switzerland. Fax: +41 22 3885601.

E-mail address: clara.james@hesge.ch (C.E. James).

meaning emerges as a function of interaction between musical formal structures, without any reference to the external world by means of words or other symbols (Koelsch, 2011).

In the context of pitch processing, a melody in major versus minor mode, or a deceptive versus a perfect cadence, convey different meanings; although hard to verbalize, then: music may express the ineffable (Jankélévitch and Abbate, 2003). Some evidence exists that cognitive mechanisms for higher order pitch processing, important for conveying semantic meaning in language, may be shared between language and music (Perrachione et al., 2013). Despite all these observations, some authors in the linguistic domain strongly deny any existence of semantics in music (Kutas and Federmeier, 2011; Pinker, 1997).

The N400 response is an event-related potential (ERP) presumed to process meaning. Initially observed in response to semantically incongruent sentence endings in language (Kutas and Hillyard, 1980), the N400 was more recently linked to incongruity of meaning in a wide variety of stimuli (Kutas and Federmeier, 2011). However, Kutas and Federmeier (2011) explicitly exclude musical stimuli from this train of results, suggesting that music does not contain semantically meaningful information.

Nevertheless Koelsch et al. (2004) could demonstrate a semantic priming effect on target word processing by means of musical primes and therefore provided evidence in favor of semantic processing in music. Primes consisted of sentences or musical passages that were either related or unrelated to a target word for their meaning. Independently of whether a musical or a language prime was used, unexpected target words provoked stronger N400 responses than target words preceded by related primes. The priming effect did not differ between language and musical primes for strength, time course or neuronal generators. This N400 response primed by musical targets was elicited by words and not by music; such a response to unexpected “extra-musical meaning” results from an association from a musical sign quality to the extra-musical world (Koelsch, 2012). Daltrozzo and Schon studied the reverse effect: the influence of visually presented words on the processing of musical excerpts (Daltrozzo and Schon, 2009). They also found a larger N400 response now in response to musical targets that were unrelated to the preceding extra-musical word context. These studies investigated people with none or low level of musical education.

In intra-musical contexts, transgression of intra-musical meaning by means of harmony transgression (thus violation of syntax) embedded in either chord series (Koelsch et al., 2000) or authentic musical stimuli (Koelsch et al., 2008), yielded an N5 response in non-musicians preceded by an ERAN (Koelsch et al., 2001).

Children and adults with and without formal musical education were able to discriminate tonal transgressions in melodies, chord progressions and complex music (Besson and Faïta, 1995; Bigand et al., 2003; Koelsch et al., 2000, 2005; Trainor and Trehub, 1994). Recently we could show that primary-school children with and without musical training perfectly ranked transgressions of musical closure as a function of their salience in monophonic children's songs but also in more complex polyphonic piano pieces (James et al., 2012). Moreover, typical ERP responses, similar to those of adults, occurred in 5- and 9-year old non-musician children in response to unattended chord series with transgressed endings (Koelsch et al., 2003). Whether these behavioral and associated ERP responses persist in primary-school children when irregularities are in-key and refined, and furthermore embedded in complex polyphonic music, has not been investigated yet.

In the current experiment we presented a series of ecologically valid short string quartets (duration approx. 10 s) in a variety of Western tonal styles, to 15 randomly assigned primary-school children of around ten years of age. The pieces ended either regularly or contained refined in-key harmonic transgressions at closure. Notably, the transgressed and thus unexpected terminal chords were perceptually just as pleasant as the regular endings (see the [Methods](#) section, musical violation

discrimination task). Meanwhile high-density electroencephalography (EEG) was recorded, that we analyzed with microstate analyses comprising all electrodes, and also with classical ERP analyses, using an electrode array deriving from a previous study (James et al., 2008).

As similar stimuli have not been presented to a child population before, our study is essentially exploratory, and cannot be placed in a full developmental context. However, we anticipated possible occurrence of early and later ERP components known to arise in response to tonal expectation violation in music in adults and children. In order to investigate whether the processing of structure and meaning in music is related to certain domain general cognitive functions, we additionally assessed measures of working memory and fluid intelligence.

Methods

Participants

Sixteen right-handed children (8 girls, 7 boys; mean age = 10.9 years, SD = 0.8) were randomly assigned among the population of a local public primary school in Geneva, Switzerland. Their parents gave written informed consent. One supplementary child was excluded from the sample because of excessive blinking during EEG recordings. The protocol was approved by the school and by the local ethics committee. Only right-handed participants were included for brain organizational reasons (Isaacs et al., 2006). Handedness was assessed with the “Edinburgh Handedness Inventory” (Oldfield, 1971; mean score 87.3 ± 16.7). According to reporting of the parents, all children possessed normal hearing and had no history of neurological illnesses. Of the 15 final participants, 10 children did not receive any extra-curricular musical training, 5 practiced musical instruments for 2.8 ± 0.5 years; one played the cello, two the flute and two the piano. Intra-curricular music lessons at public schools in the French speaking part of Switzerland are not strictly protocolled and differ strongly as a function of the teacher. Thus, some explicit knowledge of musical scales may have been provided to the non-musician children.

General procedure

All children were tested twice. At first the children accomplished a musical violation discrimination test while continuous EEG was recorded (see [Experimental procedure I: EEG](#)). On average 3 days later (SD: 5 days), the children passed 2 additional behavioral tests in the following order:

- 1) Raven's “standard progressive matrices” (RSPM; Raven et al., 1998)
- 2) “Digit span” of the Wechsler Intelligence Scale for Children (Wechsler, 2005)

We applied those tests aiming to evaluate possible relationships between short-term memory (Digit span forward), working memory (Digit span backward), fluid intelligence (RSPM) and musical violation discrimination.

Experimental procedure I: EEG

Musical violation discrimination task

We presented 78 original polyphonic expressive musical stimuli at two levels of musical transgression at musical closure: regular and harmonically¹ transgressed endings. The original stereo sound-files were converted to single track/mono stimuli in order to cancel out any left ear versus right ear differences and presented binaurally via EEG compatible headphones. The stimuli ($n = 156$) were presented in

¹ Musical harmony refers to the use of simultaneous pitches or chords.

2 different pseudo-randomized sequences, the second in reverse order of the first to prevent serial position effects. The series were presented in 4 blocks, separated by short breaks.

The musical score of an example stimulus showing both transgression levels at closure is depicted in Fig. 1; the corresponding sound-file is available in the supplementary material. Similar stimuli were used in a recent fMRI study (Oechslin et al., 2013b) that examined adults with 3 distinct levels of musical expertise (non-musicians, amateurs and experts). The latter study contained a supplementary transgression level: a very subtly transgressed ending, which was almost exclusively discriminated by professional musicians and therefore omitted here. These ecological stimuli were manufactured specifically for our experiments by a professional composer, and cover a large range of musical styles from baroque to late romanticism. They were balanced for all 24 minor and major tonalities. The compositions are string quartets with a mean duration of around 10 s, arranged with the “Sibelius” software (Avid Technology, Inc. and “Logic Pro”, Apple Inc.); real instrumental timbres (violin, viola and cello) were implemented using the “Garritan Personal Orchestra” (<http://www.garritan.com>). Regular endings consisted of a first degree or *tonic* in root position; transgressed endings consisted of the 4th degree or *sub-dominant* in first inversion. We used the first inversion (the so-called sixth chord) because this allows better voicing (vertical ordering and spacing of the pitches), making the transgressions as refined as possible, by avoiding the fourth note of the scale in the bass. Both the tonic and the subdominant chord played by a string quartet, thus widely spread over the octaves, in major and minor tonalities, in root position or first inversion, may contain the same intervals (octaves, perfect fifths, major and minor thirds and their inversions: perfect fourths, major and minor sixths) and therefore induce approximately the same degree of consonance or pleasantness.

All terminal chords were cut off at 1400 ms from onset and faded linearly over the last 150 ms. The carefully voiced transgressions all remained within the tonality (“in-key”) of the preceding context and would be suitable if the music continued. However, as endings, these chords represent transgressions according to the Western tonal rule system, only the tonic chord in root position is appropriate at closure. As we systematically conserved the melody in the upper voice between regular and transgressed endings, and also because the final chord was always in-key and carefully voiced with respect to the preceding context, these stimuli can be considered refined and thus relatively

difficult to discriminate, even more so because of the polyphonic and thus complex nature of the musical material.

Participants' appraisals consisted of expressing whether end-formulas were correct with respect to the preceding musical context, yes or no, by means of right hand button presses on a computer mouse, using the left button labeled “no” (middle finger) for incorrect endings, and the right labeled “yes” (index finger) for correct endings. We will refer to this task as the “musical violation discrimination task”. The children were instructed to withhold their response and hold still and relaxed after the onset of the final target chord, until a prompt (“please respond”) was presented on the screen, appearing 1900 ms after stimulus onset, in order to prevent contamination of the stimulus-related EEG signal with motor activity. This is why no reaction times are reported, and why the use of index and middle finger was not counterbalanced.

Based on these binary responses we computed accuracy (percent correct) of responses to regular and transgressed endings and consecutively *d*-prime values. The *d*-prime index is a statistic derived from signal detection theory (Macmillan and Creelman, 1997) and provides an index of rater sensitivity: higher *d*-prime values indicate better discrimination between regular and transgressed endings.

EEG acquisition and raw data processing

EEG was continuously recorded at 64 electrode sites (BioSemi Active-Two, V.O.F., Amsterdam, The Netherlands), equally distributed across the scalp. Data were digitized at a sampling rate of 1024 Hz in a bandwidth filter of 0–268 Hz. Prior to analysis, data were offline recomputed against average reference and band-pass filtered between 0.25 and 30 Hz within a 2nd-order Butterworth filter (–12 dB/octave roll-off). As we recorded in a classroom and not in a Faraday cage, we applied a 50 Hz notch filter in order to cancel out power-line interference. However, the BioSemi Active-Two system already annihilates the effect of interference currents in the electrode wires by using active electrodes, and is therefore ideal for field research. Average evoked potentials were computed from 100 ms before stimulus onset (i.e. the onset of the terminal chord) to 800 ms post stimulus. We then applied average referencing that consists of a spatial correction of the voltage configuration on the scalp of all electrodes at each data point by the mean voltage of all electrodes. We also applied a classical baseline correction using the 100 ms preceding stimulus onset in order to better compare the ERPs to the existing literature. We used pre-processing

The figure displays a musical score for a string quartet in D major. It consists of two parts: a regular ending (R) and a transgressed ending (T). The regular ending (R) shows the final chords for Violin I, Violin II, Alto, and Violoncelle, all resolving to the tonic (I). The transgressed ending (T) shows the final chords for Violin I, Violin II, Alto, and Violoncelle, all resolving to the first inversion of the fourth degree or subdominant (IV⁶). The score includes dynamic markings such as *f*, *mf*, *cresc.*, *pizz.*, and *arco*. The regular ending is marked with an asterisk (*) and the transgressed ending is marked with a circled asterisk (*).

Fig. 1. This figure displays the score of an example musical stimulus, here in D major. All musical pieces were presented in 2 versions: with a regular ending (R, piece ends on the tonic, “I”), or a harmonic transgression at closure (T, the piece ends on the first inversion of the fourth degree or subdominant “IV⁶”). Corresponding audio files are provided in the supplementary material (respectively D_Maj_R.wav, D_Maj_T.wav).

methods as described in Brunet et al. and Michel and Brandeis (Brunet et al., 2011; Michel and Brandeis, 2009): in addition to an automated threshold rejection criterion of 100 μ V, all epochs were visually inspected for oculomotor and other artifacts (movement, muscular tension, loss of electrode contact following discrete movement or perspiration). Only artifact-free epochs were retained with an average of 63.9 ± 6 epochs per condition per participant.

Channels exhibiting substantial noise were interpolated with a 3D spherical spline interpolation accounting for the real geometry of the head (Brunet et al., 2011; Perrin et al., 1989). These interpolation methods are recommended for high density ERP source analysis (Brunet et al., 2011; Michel and Brandeis, 2009).

Procedure of ERP analyses

The ERPs were analyzed in 4 consecutive stages.

Stage 1: ERP waveform analyses. We computed average ERPs in response to both experimental conditions and also ERP difference waves in an 8-electrode site array divided into two zones, frontal (F1, F2, Fz and Afz), and parietal (P1, P2, Pz and POz). To compute difference waves, ERP responses of regular endings were subtracted from ERP responses to transgressed endings.

The choice of these electrodes was based on an ERP response of non-musician adults revealed by a previous study (James et al., 2008), using similar harmonic transgressions in piano pieces, and also on visual inspection of the current data, because we did use a 64 electrode recording here and a 128 electrode setting in James et al. (2008).

We conducted a repeated-measures ANOVA on the mean amplitude of ERPs over the time period 350–450 ms after stimulus onset with the factors Condition (2, within) \times Zone (2, within).

This time period corresponds with the robust peak latency of the N400 component (Kutas and Federmeier, 2011) and also comprises the period of occurrence of the ERP responses of adult non-musicians mentioned above (James et al., 2008). The peak latency of the Central Posterior Negativity (CPN) observed here for the 8 chosen electrodes was 407.9 ± 3.6 ms. For the difference waves (subtraction of ERP responses to regular endings from ERP responses to transgressed endings) of these 8 electrodes, the peak latency was 408.9 ± 2.6 ms.

Stage 2: spatiotemporal ERP analyses I: topographic dissimilarity analysis.

In a second stage of the analysis we computed a measure of topographic dissimilarity over time that allowed determining whether the observed differences between the ERPs of the two experimental conditions in stage 1 were due to topographical changes of the whole scalp potential configuration or to local amplitude changes only. According to physical laws, only if the whole scalp potential configuration changes significantly over time, differences in underlying active generators in the brain can be assumed (Vaughan, 1982).

In order to check for such topographic modulations, we performed statistical comparisons between the two experimental conditions of the mean normalized ERP topographies at each time point using a nonparametric bootstrapping method on the global map dissimilarity values (Kondakor et al., 1997; Michel et al., 2004b; Murray et al., 2006, 2008; Srebro, 1996). Global dissimilarity is a single measure expressing the topographic difference between 2 scalp voltage topographies, normalized to unitary strength by dividing all amplitudes by the GFP at each time point (Lehmann and Skrandies, 1980; Murray et al., 2008). For this analysis we only retained periods lasting for at least 30 consecutive ms during which this topographic test exceeded a 0.05 alpha criterion.

Stage 3: spatiotemporal ERP analyses II: microstate analysis & statistical fitting procedure. Classical ERP waveform analysis focuses on explicit components defined in time and space. This analysis allows detecting differences between populations or conditions over time.

In microstate analysis we study the scalp voltage topographies over time of all recorded electrodes in one unified analysis. These microstates reflect the sum of all neuronal populations active at a given moment in time and manifest as discrete segments of electrical stability lasting tens to hundreds of milliseconds separated by brief intervening intervals (Brunet et al., 2011; Murray et al., 2008; Pascual-Marqui et al., 1995). Lehmann and colleagues (1987) conceptualized these stable periods as *microstates* of information processing, basic psychophysiological units of cognition.

The principal aim of the microstate analysis is to look for changes in configuration of the potential field over time that directly informs when different generators are active in the brain.

Microstate analysis thus provides intrinsically different information as compared to classical ERP analysis, and the 2 analyses can be considered complementary.

We applied a hierarchical cluster analysis (CarTool software; <http://brainmapping.unige.ch/cartool.htm>) to reduce our grand-average ERP topography time series of both experimental conditions into an optimal number of microstates or map topographies. The optimal number is determined based on cross validation criteria that minimize the residual variance, and the Krzanowski–Lai criterion (Brunet et al., 2011; Michel and Brandeis, 2009; Michel et al., 2004a; Murray et al., 2008; Pascual-Marqui et al., 1995; Tibshirani and Walther, 2005). The resulting series of microstates over time can be considered an a priori hypothesis that has to be statistically tested. This statistical testing is done by fitting the microstates to the time series of each individual participant. For that purpose spatial correlation coefficients are computed for each time-point between the microstates identified by the cluster analysis in the grand-average ERPs and the ERPs of each participant for each condition. This procedure yields measures of microstate presence or “microstate duration” (expressed in ms) for each participant and condition that can be submitted to statistical testing, thus “fitting” the microstates to the actual data.

The distributions of the microstate durations showed some asymmetry. But no outliers were present, so we used repeated measures ANOVAs to compare microstate durations between the experimental conditions. As no reliable tests exist to verify normality with small sample sizes (Razali and Wah, 2011), we verified these results by means of non-parametric contrasts (Wilcoxon signed-rank test).

Stage 4: statistical analysis of ERP sources: estimation & comparison of neuronal generators.

Because the inverse problem is ill-posed with scalp EEG, the inverse solution to estimate neuronal generators must be stabilized via a priori constraints (Grech et al., 2008; Michel et al., 2004a). Statistically verified microstate series provide such a priori constraints and consequently reduce the dimensionality of the inverse problem. We therefore limited source estimations to time periods during which spatiotemporal analysis of scalp ERP demonstrated significantly different microstates/map topographies between stimulus types. As already stated in the preceding section, changes of scalp voltage configurations are indicative of changes in the underlying generator configuration (Michel et al., 2004a; James et al., 2008; Murray et al., 2008; Vaughan, 1982).

As a first step we estimated the intracranial current distribution at each moment in time for the evoked potential of each subject in each condition with a depth-weighted minimum norm (WMN) distributed linear inverse solution (Hamalainen and Ilmoniemi, 1994; Michel et al., 2004a). The current distribution was calculated within the gray matter of the average brain provided by the Montreal Neurological Institute. A discrete grid of 3005 solution points was regularly distributed within this volume. After applying a homogeneous transformation operation to the volume that rendered it to the best fitting sphere (SMAC model; Spinelli et al., 2000), a 3-shell spherical head model was used to calculate the lead field for the 64 electrodes and the inverse solution based on the weighted minimum norm (WMN) constraint. As

a second step, Bonferroni corrected two-tailed t-tests were applied to compare mean current density values over each selected time period between both experimental conditions; a procedure of statistical parametric mapping (SPM). In distributed EEG/MEG inverse solutions, the number of independent variables used for the Bonferroni corrections corresponds to the number of recording sensors on the scalp (Grave de Peralta Menendez et al., 2004; Michel et al., 2004a; Murray et al., 2008). This procedure allows detecting brain areas in which enhanced activity occurs for either one of the experimental conditions (or groups) and also cancels out residual random noise present in both conditions. These methods have been successfully used in many domains recently (Britz et al., 2009, 2014; James et al., 2008; Plomp et al., 2013; Rihs et al., 2013). However, source imaging from grand-averaged EEG data has limited accuracy (Plomp et al., 2010) and the areas we depicted as generators should therefore be interpreted as coarse estimates.

Experimental procedure II: additional behavioral measures

"Digit span" of the Wechsler Intelligence Scale for Children (Wechsler, 2005)

The Memory for Digit Span assessment, a component of the Wechsler Intelligence Scales for Children—Revised (WISC-R), is a measure of short-term memory for children aged seven and over (Wechsler, 1974). During these tests children are orally given sequences of numbers and asked to repeat them, either as heard (Digit span forward) or in reverse order (Digit span backward). Digit span forward (DSF) and Digit span backward (DSB) assess distinct but interdependent cognitive functions (Grégoire, 2009). DSF primarily evaluates short-term auditory memory whereas DSB measures the child's ability to manipulate verbal information while in temporary storage thus working memory capacity.

Raven's "standard progressive matrices" (Raven et al., 1998)

We used Raven's "standard progressive matrices" (RSPM) (Raven et al., 1998) in a time-limited version (20 min) to assess a measure of fluid/general intelligence. Fluid intelligence is the capacity to reason and solve new problems without relying on previously acquired knowledge and skills. The test consists of 60 visuo-spatial problems that become increasingly difficult. The items require inferring a rule relative to a collection of visuo-spatial patterns presented in a matrix. The subject has to identify the last missing element of the pattern from a certain number of choices provided (6 or 8 options), based on the previously inferred rule.

Results

Behavioral results

Musical violation discrimination task

On average, the children discriminated the regular endings ("R"; $82.05 \pm 17.42\%$ (SD)) better than the harmonic transgressions ("T"; $67.27 \pm 18.57\%$; $t_{14} = 2.86$, $p = 0.01$). They discriminated both R ($t_{14} = 7.12$, $p < 0.01$) and T ($t_{14} = 3.60$, $p < 0.01$) well above chance; mean d-prime values were 1.74 ± 1.24 points. No significant difference existed between the d-prime scores of musician ($n = 5$) and non-musician ($n = 10$) children ($t_{13} = 1.26$, $p = 0.23$).

Additional behavioral tests

RSPM yielded raw scores of 44.60 ± 4.10 . As no test norms exist for the limited time version for French 10 year-olds, we could not compare our sample to the general population.

The Memory for Digit Span assessment yielded mean scores (raw data) for recall of numbers for recall of numbers in direct order (DSF) of 8.80 ± 1.78 versus 7.27 ± 2.28 for those in inverted order (DSB). As typically observed, scores for direct order were superior compared to those in inverted order ($t_{14} = 2.20$, $p = 0.04$). The scores were within the norms for this age group.

Correlation analysis of behavioral results

We computed Pearson correlations between the d-prime scores resulting from the musical violation discrimination task and the three other behavioral measures: DSF, DSB and RSPM scores. Only the linear relationship between DSB and d-prime ($r = 0.53$, $p = 0.04$) was statistically significant, no significant correlation existed between DSF and d-prime ($r = 0.33$, $p = 0.24$) or between RSPM and d-prime ($r = 0.37$, $p = 0.17$).

ERP results

Stage 1: ERP waveform analyses

We computed a repeated measures ANOVA on mean amplitude of ERPs in a 350–450 ms window for the 8 selected electrodes (see the Methods section), with the factors Condition (R vs. T; within) and Zone (frontal vs. parietal; within). The analysis revealed significant interaction between both factors ($F_{1,14} = 17.45$, $p < 0.01$); no main effects of Condition and Zone were observed, because in the frontal and parietal zones effects were opposite for the 2 conditions, thus resulting in means close to zero. Contrasts for all electrodes between the 2 conditions, corrected for multiple comparisons (Bonferroni), disclosed significant effects of Condition for each separate electrode of the array (Fig. 2) within the 350–450 ms period.

We also computed difference waves (T-R) that showed pronounced deviations from 0 μ V peaking around 400 ms after stimulus onset, with strong positive values for the 4 frontal electrodes and strong negative values for the 4 parietal ones in a remarkably symmetric configuration (Fig. 3).

Correlation between d-prime and CPN component. To investigate the relationship between electrophysiological data and behavior, we computed Pearson correlations between d-prime scores and the CPN component. We performed correlation analyses on d-prime scores and mean amplitude of the ERP difference waves (350–450 ms) for the frontal zone (F1, Fz, F2 and Afz) and the parietal zone (P1, Pz, P2 and POz) of the 8-electrode array. We chose to use ERP difference waves because, like the d-prime score, they express a relationship between responses to transgressed and regular endings. For both zones, mean amplitude and d-prime scores were significantly correlated (frontal zone: $r = 0.62$, $p = 0.01$; parietal zone: $r = -0.57$; $p = 0.03$). So higher amplitudes of difference waves for this component related linearly to higher d-prime scores that exhibit higher rater sensitivity to the harmonic transgressions.

Stage 2: spatiotemporal ERP analyses I: topographic dissimilarity analysis

The topographic dissimilarity analysis (Fig. 4a) revealed a sustained difference between the 2 experimental conditions from 321 to 639 ms, and a quasi-sustained difference from 271 to 755 ms (Fig. 4a; depicted values correspond to $p < 0.05$, duration > 30 ms; periods of significant difference: 271–312 ms; 321–639 ms; 654–690 ms; 714–755 ms).

Stage 3: spatiotemporal ERP analyses II: microstate analysis & single-subject "fitting"

The hierarchical cluster analysis yielded 7 distinct microstates that optimally represent the data for both conditions and all participants (Figs. 4b & c); this microstate series explained 92.22% of the variance in the ERP data during the time period 0–800 ms. Visual inspection of the occurrence of the microstates (M) over time disclosed several differences in the grand means between responses to R and T.

An early difference, exclusive occurrence of M4 for R versus M1 for T (126–255 ms), could be observed in the grand means. After ≈ 275 ms after stimulus onset M2, M6 and M7 occurred exclusively in condition T, M3 and M5 in condition R. Observing the scalp voltage configuration, M2 can clearly be associated with the CPN component observed in the ERP waveform analyses.

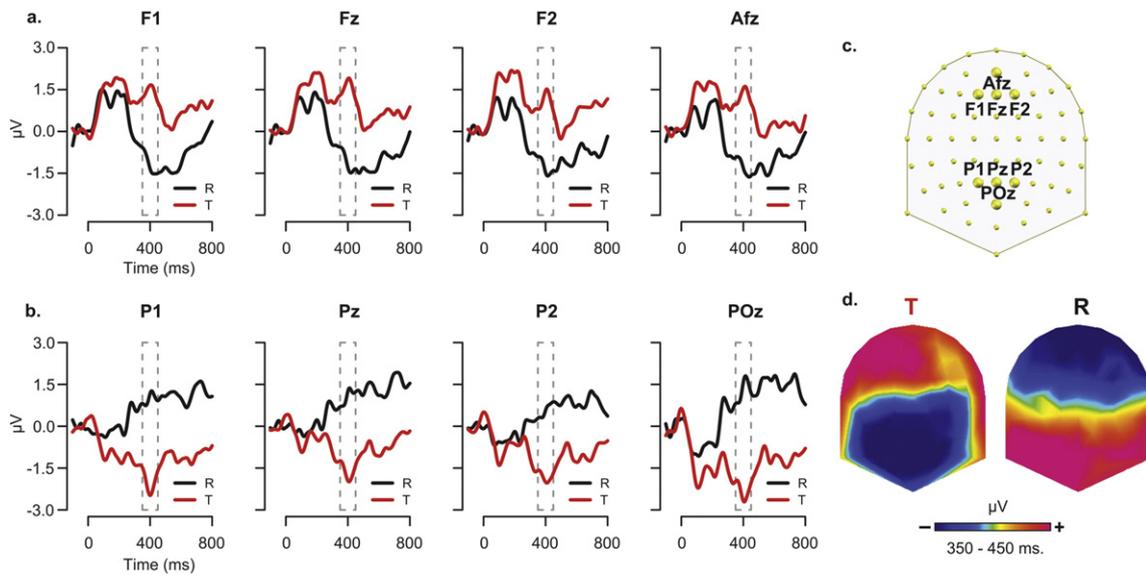


Fig. 2. Grand-average ERP waveforms at 8 electrode sites at (a) 4 frontal and (b) 4 parietal electrode sites. ERPs elicited by regular terminal chords (R) are plotted in black, and by transgressed ones (T) in red. The time interval used for the statistical analysis of the CPN component is indicated by gray-dashed brackets (350–450 ms). (c) Scalp positions of the 8 electrodes are highlighted within the full 64-electrode setting (view from above, nasion on top). (d) Mean whole head voltage configurations are shown for both conditions R and T over the 350–450 ms period.

We then applied the statistical fitting procedure on the variable “microstate duration” (in ms; see the [Methods](#) section) over 4 consecutive time periods during which 2 different microstates occurred simultaneously in both conditions in the grand means. This fitting is necessary to verify the statistical significance of the observed differences in the grand means. We will report only interaction effects (main effects of condition will not occur because the total amount of ms is always equal for both conditions over each time period). Main effects of microstate that do not take into account the factor Condition are not relevant in the context of the present study. As the first period did not yield any significant results, we labeled it Period 0.

Period 0 (126–255 ms) results from the following boundaries: 126 ms, appearance of M4; 255 ms, extinction of M4. A repeated measures ANOVA (Condition (R vs T) \times Microstate (M1 vs M4)) did not show significant interaction Condition \times Microstate ($F_{1,14} = 2.94, p = 0.11$). This is not surprising, as M1 and M4 display a very similar voltage configuration. Moreover, the topographic dissimilarity analysis did not reveal significant differences either over this period ([Fig. 4a](#)).

Period 1 ([Fig. 4d1](#)) 301–477 ms results from the following boundaries: 301 ms, appearance of M3 for R, 477 ms, extinction of M2

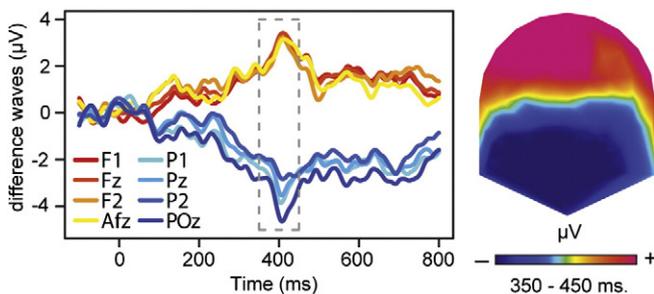


Fig. 3. Grand-average ERP difference waveforms (T-R) at 8 electrode sites (cf. [Fig. 2c](#)), 4 frontal (warm colors) and 4 parietal (cold colors). Note the marked symmetric difference from zero peaking around 400 ms. On the right, the mean whole head voltage configuration of the difference waves is depicted over the 350–450 ms period, corresponding to the gray dashed box.

for T. A repeated measures ANOVA (Condition (R vs T) \times Microstate (M2 vs M3)) showed significant interaction Condition \times Microstate ($F_{1,14} = 17.39, p = 0.01$). M2 appeared on average 37.89 ± 53.44 ms in condition R and 136.98 ± 64.53 ms in condition T. M3 appeared on average 138.87 ± 53.44 ms for R and 39.78 ± 64.53 ms for T. The observed dominance of M2 for T ($z = 3.01, p < 0.01$) and M3 for R ($z = 3.01, p < 0.01$) could be confirmed by means of Wilcoxon signed-rank tests. In order to ensure that this main result, consistent with the ERP waveform results of the CPN component, was not yielded mainly by the 5 musician children, we repeated this test including exclusively non-musician children ($n = 10$); an identically oriented statistically significant interaction was obtained between Condition and Microstate ($F_{1,9} = 9.94, p = 0.01$) confirming that this effect manifested in all children.

Period 2 ([Fig. 4d2](#)) 477–603 ms results from the following boundaries: 477 and 603 ms, respectively appearance and distinction of M6 for T. A repeated measures ANOVA (Condition (R vs T) \times Microstate (M3 vs M6)) showed significant interaction Condition \times Microstate ($F_{1,14} = 8.95, p = 0.01$). M3 appeared on average 106.84 ± 35.51 ms for R and 50.65 ± 49.77 ms for T. M6 appeared on average 20.12 ± 35.51 ms for R and 76.30 ± 49.77 ms for T. The observed dominance of M6 for T ($z = 2.54, p = 0.01$) and M3 for R ($z = 2.54, p = 0.01$) could be confirmed by means of Wilcoxon signed-rank tests.

Period 3 ([Fig. 4d3](#)) 644–800 ms results from the following boundaries: 644 ms, appearance of M5 for R, 800 ms, end of analysis. A repeated measures ANOVA (Condition (R vs T) \times Microstate (M5 vs M7)) showed significant interaction Condition \times Microstate ($F_{1,14} = 8.95, p = 0.01$). M5 appeared on average 120.44 ± 40.64 ms for R and 63.48 ± 57.43 ms for T. M7 appeared on average 35.81 ± 40.64 ms for R and 92.77 ± 57.43 ms for T. The observed dominance of M7 for T ($z = 2.29, p = 0.03$) and M5 for R ($z = 2.29, p = 0.03$) could be confirmed by means of Wilcoxon signed-rank tests.

Nota bene: Effects for Wilcoxon signed-rank tests for both microstates within each time period yielded always identical z and p -values, because the total amount of ms is the same.

In conclusion, over all 3 periods M3 and M5 are prevalent for R, characterized by posterior positivity, against prevalence of M2, M6 and M7 for T, characterized by posterior negativity.

Stage 4: statistical analysis of ERP sources: estimation & comparison of neuronal generators

We used the results from stage 3 (microstate analysis) of the ERP analyses to determine periods of time with potential distinct underlying

generators in the brain for both experimental conditions (R and T). For each of the 3 periods for which the fitting procedure disclosed a significant interaction between microstate incidence of R and T (see Fig. 4d.), we computed statistical comparisons (statistical parametric mapping),

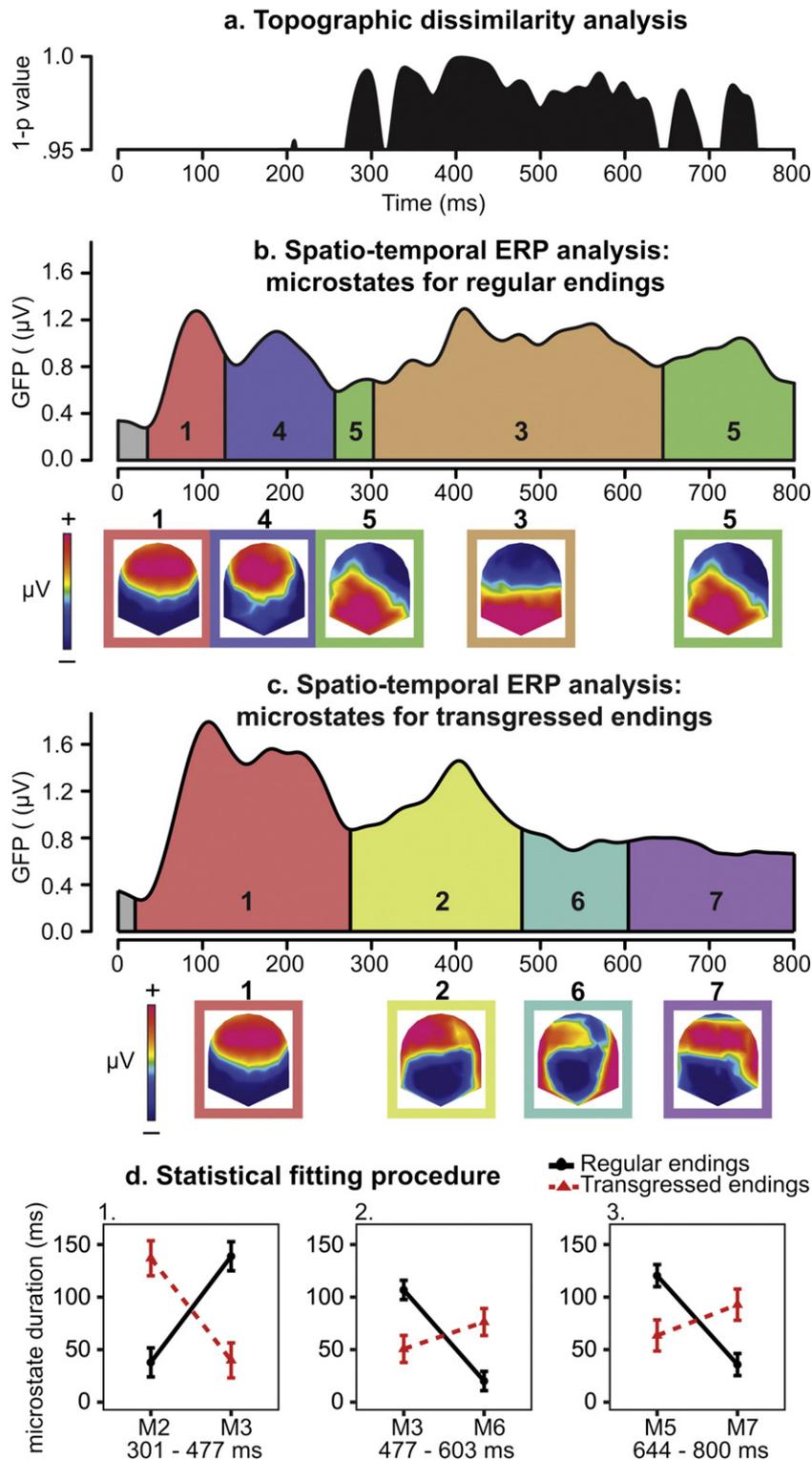


Fig. 4. Spatiotemporal ERP analyses. (a) Topographic dissimilarity analysis, 1-p values are plotted over time. (b, c) A hierarchical cluster analysis yielded 7 distinct microstates that optimally represent the data of both conditions of all participants. (b) Microstates for regular endings, top row: the colored segments under the global field power (GFP) curves represent the time periods during which each of these microstates was most represented in the group data; bottom row: scalp voltage configurations of the microstates (view from above, nasion on top), framed in corresponding color-code. (c) Idem for transgressed endings. (d1–3) Fitting of microstates (M) for 3 consecutive time periods during which significant interaction Condition \times Microstate occurred. On the y-axis mean duration (ms) of microstates is provided. Vertical bars depict standard errors.

through Bonferroni corrected two-tailed *t*-tests,² between mean current density values (expressed $\mu\text{A}/\text{m}^3$) in response to R and T over these 3 time periods: 301–477 ms, 477–603 ms and 644–800 ms after stimulus onset.

The *t*-tests on mean current density over the full 301–477 ms window did not yield significant results. Lowering the threshold from Bonferroni corrected to a more clement $p < 0.005$ we found stronger activations for condition T as compared to R in bilateral pre-motor areas in the medial frontal gyrus and paracentral lobule and in a discrete left auditory area with solution points in the superior and middle temporal gyrus. In order to conserve a Bonferroni corrected threshold, we narrowed the time window around the marked GFP peak of the CPN response, enhancing the signal-to-noise ratio that is known to increase with GFP strength (Lehmann et al., 2005). Stable map configuration and underlying cerebral networks are presumed during micro-states (Koenig et al., 2014), and map stability is at its maximum during the GFP peak. Computing sources at the GFP peak using SPM was recently successfully used to show deficits of the ventral stream in visual processing in schizophrenic patients (Plomp et al., 2013). In a time window from 375 to 425 ms, we found Bonferroni corrected significant differences in large clusters, with higher current density in response to condition T in a network of brain areas with peak values in the left hemisphere. The network comprised voxels in the medial frontal gyrus ($-3 -21 52$; BA6; $t = 5.03$, $p = 0.00019$), in the paracentral lobule (peak value Talairach $-9 -41 54$; BA5; $t = 6.48$, $p = 0.00002$ & $-3 -34 66$; BA4; $t = 4.42$, $p = 0.00058$), in the precuneus (peak value Talairach $-3 -47 54$; BA7; $t = 4.60$, $p = 0.00041$) and in the mid-posterior cingulate gyrus (peak value Talairach $-3 -28 42$; BA31; $t = 4.33$, $p = 0.00069$; see Fig. 5a).

The *t*-tests on mean current density over the following 477–603 ms period, displayed significant differences in a network in the right hemisphere, again with higher values for condition T, in voxels of the postcentral gyrus (peak value Talairach $54 -11 14$; white matter; $t = 5.22$, $p = 0.00013$), of the superior temporal gyrus (peak value Talairach $53 -11 9$, BA22; $t = 5.22$, $p = 0.00013$; peak value Talairach $55 -27 14$, BA 42; $t = 4.38$, $p = 0.00063$) and in a sub-lobar part of the mid-posterior insula (peak value Talairach $47 -12 4$; BA22; $t = 4.50$, $p = 0.00050$; see Fig. 5b). With a more clement threshold of $p < 0.005$ without Bonferroni correction, also voxels in the anterior insula (BA 13) reached significance.

In the last time period, extending to the end of our window of analysis, 644–800 ms, differences arose in a single left hemisphere cluster located in the middle temporal gyrus (peak value Talairach $-61 -38 -8$; BA21; $t = 5.24$, $p = 0.00013$; see Fig. 5c).

Discussion

The key finding of this study is the occurrence of a specific neural correlate, a centro-posterior negativity (CPN), in response to musical violation expectation in primary-school children of approximately 10 years of age. The stimuli that yielded this response were refined in-key transgressions of harmony embedded in expressive string-quartets; such transgressions may be considered syntactical as they violate structural rules within the Western tonal system. Probably due to the subtle nature of the transgressions and the complex musical context, typical earlier and later ERP components, often shown in response to musical syntactical incongruity, did not occur here. The latency of the CPN was closely centered on 400 ms, like the classical N400 observed in response to semantically incongruent sentence endings in language (Friederici et al., 1999; Lau et al., 2008). Rater sensitivity (*d*-prime) for transgressed endings was positively correlated with the strength of this CPN and also to a measure of working memory

capacity. Statistical parametric mapping (SPM) of estimated sources, comparing responses to transgressed and regular endings showed enhanced brain responses to transgression in left pre-motor, mid-posterior cingulate and superior parietal cortex. These brain areas may be linked to information integration processing. Altogether, these results induce questioning on the level of differentiation of semantic and syntactic processing in a complex intra-musical context in children.

Behavioral results

Interestingly, the randomly assigned child population discriminated the refined transgressions well above chance, but regular endings were better rated, showing that this is an easier task. This observation confirms earlier findings in musically naïve adults (James et al., 2008; Koelsch et al., 2007).

Event-related potentials in response to tonal expectation violation in music

In an extensive review article on the N400 (Kutas and Federmeier, 2011) written by the same author that first described it in 1980 (Kutas and Hillyard, 1980) the authors state “However, N400-like activity was not observed in response to unexpected events in other structured domains such as music, be they the frequency of a note violating a musical scale sequence or a familiar melody” (p. 624; Kutas and Federmeier, 2011).

N400-like responses arose in response to written, spoken, signed words and pseudowords, and outside of the language domain in response to pictures, film, sounds and mathematical symbols (Kutas and Federmeier, 2011; Niedeggen et al., 1999), even odors (Castle et al., 2000). Hence Kutas and Federmeier (2011) reason that the N400 rather conveys processing of meaning, independently of the linguistic dimension. Defined more precisely, the N400 reflects *semantic integration of a word or other percept within its working context*. Because it is harder to integrate an element in an implausible continuation (Lau et al., 2008), incongruous elements provoke the N400 response. This integration process comprises a semantic memory function that compares an element with its current context, and also to prior knowledge. Working memory function enables such comparison to the current context.

Altogether there is no theoretical reason why an intra-musical context would not be able to evoke an N400-like response. We consider that in order to evoke such a response, specific musical material (prime and target) should be presented. Our expressive stimuli, which are complete mini-compositions of approx. 10 s, written by a professional composer, allow a comprehensive build-up of musical expectancy. Musical expectancy yields a process of integration and evaluates the congruence of musical elements within their current context, and also as a function of prior knowledge. In the case of non-musicians this knowledge of the tonal system may be acquired by implicit statistical learning (Schon and Francois, 2011), following mere exposure. The refined harmony transgressions at closure applied in the current study transgress structure but possibly also non-referential meaning. Such transgressions leave the listener puzzled, because these chords would be appropriate if the music continued, but not as an ending. Moreover, the fact that 78 different compositions were presented, in all 24 minor and major tonalities, allows drawing more generalized conclusions on music processing in a Western tonal context; the degree of variability excludes any memory or habituation interferences.

The intra-musical CPN component observed here manifests some difference with the standard N400 observed in adults in response to sentential semantic incongruity (Lau et al., 2008), or conveyed by incongruent linguistic and musical primes in extra-musical contexts (Koelsch et al., 2004). The component described in the current experiment displays a strong negative voltage as such, whereas the standard N400 reflecting semantic incongruity is in general only negative as a difference wave, i.e. the result of subtracting the response to the

² The Bonferroni correction corresponds to $\alpha = 0.05/64$ (number of electrodes) = 0.00078; for the sake of transparency we will report non-corrected values for which $p < 0.00078$.

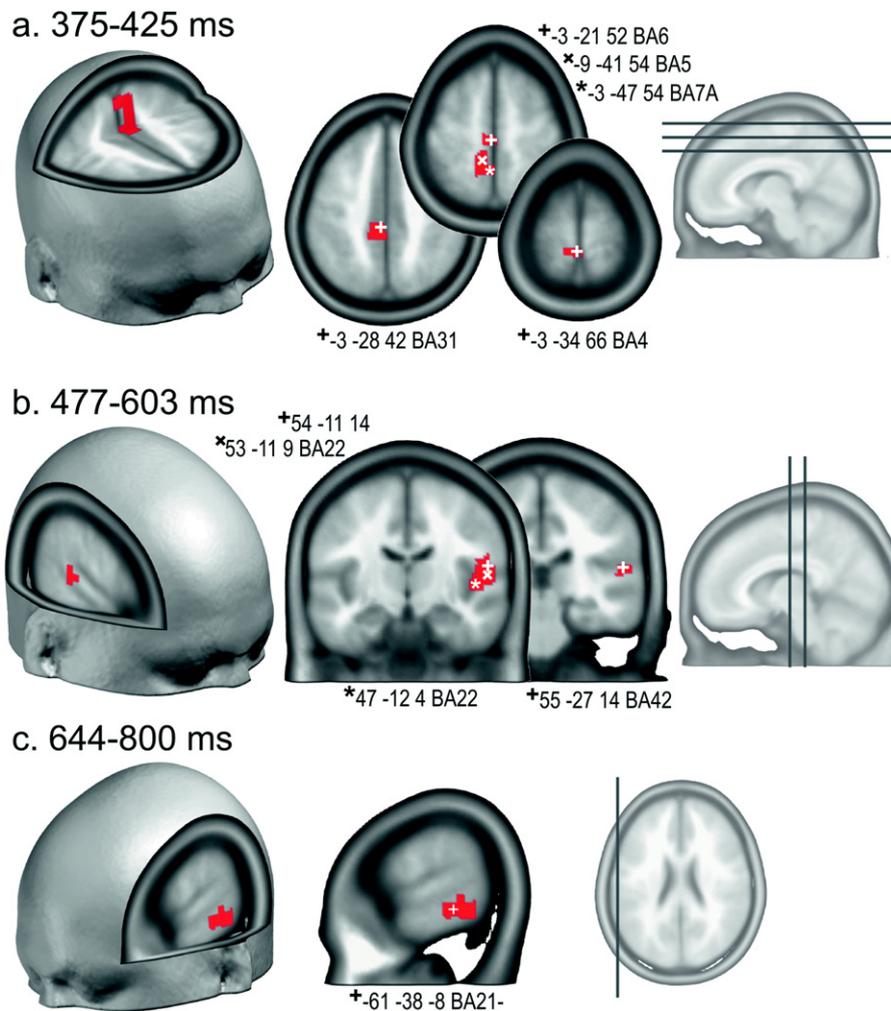


Fig. 5. Statistical parametric mapping (SPM) of estimated sources, comparing responses to transgressed and regular endings for 3 consecutive time-periods resulting from the microstate analyses (Stage 3 of ERP analyses). (a) 375–425 ms after stimulus onset (onset of the terminal chord). (b) 477–603 ms after stimulus onset. (c) 644–800 ms after stimulus onset. The highlighted brain areas display voxels with enhanced activation in response to the transgressed endings over the 3 periods. The estimated sources used for the SPM were computed via a distributed linear inverse solution (WMN). Talairach coordinates are provided, corresponding to the position of the superimposed white signs that represent the peak voxels. Corresponding Brodmann areas (BA) are provided when available.

congruent target from the response to the transgressed one. Part of this dissimilarity may be explained by our child population, in which the N400 showed increased amplitude compared to adults in language contexts (Benau et al., 2011).

The robust component observed here in response to transgression in complex music is clearly not an N5 response that has a more diffuse peak, occurring often well after 500 ms (see for example Koelsch et al., 2000). Then, although our transgressions undeniably violate structure, no ERAN or other early negativity was observed (see the following section) which typically precedes the N5 in response to syntactic violation. Interestingly, an exception to the rule that the N5 is preceded by an ERAN was also observed in another study that used in-key subdominant chords as intra-musical incongruent targets at musical closure (Poulin-Charronnat et al., 2006). So we may wonder how typical the N5 and ERAN are for syntactic and the N400 for semantic violation, and also, how distinct semantic and syntactic processing are in an intra-musical context? We are not the first to consider this query (Aramaki et al., 2010; Besson et al., 2011; Featherstone et al., 2013).

The fact that our subjects are children of around 10 years of age should also be considered. In 5- and 9-year-olds (Koelsch et al., 2003), an ERAN as well as an N5 response to harmonic incongruity in chord series could be found (Jentschke and Koelsch, 2009). Other authors also showed early negative responses (around or before 300 ms after

stimulus onset) to small pitch incongruities (“out of tune” tones) at the end of simple melodies in musician children (Magne et al., 2006; Moreno et al., 2009). In an fMRI study, activation patterns following harmonic incongruent chords in 10-year-old children were very similar to those of adults (Koelsch et al., 2005); laterality effects (more involvement of the left hemisphere in children) suggested that children process music and language more similarly than adults. At ten or eleven years of age, ERAN responses were indistinguishable from adult responses (Jentschke and Koelsch, 2009). Comparing 10-year-old children and adults for N400 effects in a traditional language paradigm using moderately and strongly incongruent sentence endings, did not show latency differences for the N400 (Benau et al., 2011); but amplitude of the N400 was enhanced in children compared to adults. Musical abilities following implicit learning seem to mature early, like those for language, between 9 and 11 years of age (Brandt et al., 2012; Pulli et al., 2008). Altogether we expected similar responses in our child population as we would in adults, with possibly stronger amplitudes, and maybe responses that are more similar to those in language than in adults, resulting in different topographies.

Now why did we observe a CPN response as the earliest distinct electrophysiological reaction to the transgressions and not an ERAN that did occur in children of the same age as our participants in response to music-theoretically similar violations? The ERAN manifested in non-musician and musician adults in response to relatively salient

transgressions (Neapolitan chords), but also in response to a supertonic chord containing no out-of-key notes (Koelsch et al., 2001, 2007). This early negative component (peak latency \approx 200 ms in adults) is a bottom-up, learned but automatized pre-attentive response to relatively salient incongruity that violates syntax. A recent neurocognitive model of music perception (Koelsch, 2011) proposed that the ERAN is involved in *structure building*. Later responses then rather reflect top-down processing, the P3a processing of novelty or surprise, the P3b context-based memory updating (Polich and Criado, 2006), with both P3 components appearing in response to rare stimuli that are task relevant and therefore only present in overt tasks, when subjects deliberately judge the transgressions. The N5 that typically follows the ERAN also when the task is covert (in many studies subjects discriminate timbre differences ignoring harmony), represents harmonic integration and meaning according to the authors that observed it (Koelsch et al., 2008; Poulin-Charronnat et al., 2006). A non-expected chord will be harder to integrate. The N5 occurred only in musical contexts, whereas the N400 is not music specific (Koelsch, 2011). Another important difference is that the N5 does not show clear peak latency like the N400, but rather a diffuse negative deflection, and, as its name reveals, has a longer latency than the N400, that may occur well beyond 500 ms after stimulus presentation.

We suggest that the refinedness or ambiguousness of the transgressions used here in combination with the rich expressive and polyphone musical context prevented typical ERP responses to musical syntactical incongruity and rarity to occur (ERAN, P3a and P3b). Exactly the same syntactical transgressions, music-theoretically speaking, may induce very different behavioral and ERP responses as a function of the context they occur in. The stimuli used in the current experiment induced processing that may be considered syntactic, but also semantic in nature. In a preceding study, we observed a similar microstate in non-musician adults in response to in-key incongruous chords (James et al., 2008, Fig. 3, microstate 5), although it did not show a sharp peak around 400 ms as in the current experiment. Yet in that study the focus was on comparing experts to non-musicians, therefore the comparison between non-musician responses to transgressed and regular endings was not directly made. By all means, the high-level experts in that study did not show this microstate in response to transgression. Although 5 musically experienced children were present in our sample here, their rater sensitivity was not different from the non-musician children, and in the 10 non-musician children the CPN clearly manifested. The expert pianists in James et al. (2008) had practiced their instrument on average for 20 years, and peak value of training was over 6 h per day. Therefore it seems legitimate to consider our participant group here as rather inexperienced listeners, and the CPN as a “non-musician” response.

Microstates in response to transgressed endings following the CPN component in time continued to display posterior negativity (microstates 6 and 7), as if the wondering about the sense of the musical phrase went on. Surprisingly, we rather observed a later posterior positivity in response to regular endings (microstates 3 and 5). Again this converges with previous findings with similar stimuli (James et al., 2008; Fig. 4 microstate 7). Some resemblance may exist between these later positivities and the so called “music closure positive shift” (Knosche et al., 2005; Neuhaus et al., 2006). After all, only regular endings provide a satisfactory closure.

Generators

While centering narrowly on the peak of the CPN, SPM of brain sources revealed enhanced brain activation for transgressed endings in the left medial frontal gyrus, paracentral and mid-posterior cingulate areas and in the precuneus. This contrasts with the literature on underlying brain sources of semantic priming (Lau et al., 2008) and N400 (Koelsch et al., 2004) that suggests temporal and inferior frontal sources, although the results are somewhat divergent.

But the present study applied SPM on distributed inverse solutions over a temporally concise period close to the peak latency of the N400. This contrasts to previous studies reviewed by Lau et al. (2008): either windows of analysis were much larger (around 200 ms), or target brain regions were identified by fMRI activations, supported by low temporal resolution (Steinbeis and Koelsch, 2008).

Although results of the computation of distributed inverse solutions should be treated with caution (see *Methods*), the strong pre-motor activations observed here that distinguish processing of transgression from processing of regular endings are plausible within the context of integration processing.

Motor activations accompanying musical perception are common in musicians and non-musicians (James et al., 2008; Sammler et al., 2013; Zatorre et al., 2007). Premotor areas may be involved in some form of covert mimetic representation of the musical material (Cox, 2011; Glenberg, 2010; Herholz et al., 2012), which supports integration processing, and is thus enhanced in the presence of incongruent information. Specifically the supplementary motor area (SMA) may initiate this simulation process, acting as a forward model that is the fruit of previous – possibly implicit – learning. Interestingly, this error prediction function also occurs in the context of complex music that the listener may not at all be able to produce (Schubotz, 2007). Such SMA activation in response to incongruity has been observed in musical, linguistic and sensorimotor contexts (Herholz et al., 2012; James et al., 2008; Meyer et al., 2011; Ye et al., 2014). Finally, posterior cingulate activation could be associated previously to an N400 in response to semantic incongruity (Frishkoff et al., 2004) – and the precuneus is a known hub for pitch processing (Platel et al., 1997; Satoh et al., 2001).

Classically reported sources for semantic priming/the N400 (Kutas and Federmeier, 2011; Lau et al., 2008): middle and superior temporal lobe and inferior frontal areas were also activated more strongly in response to transgressed endings in the current experiment, but later in time. The right superior temporal cortex and insula responded to transgression in period 2 (477–603 ms), the left middle temporal gyrus in period 3 (645–800 ms). The middle and superior temporal gyrus, together with the inferior frontal cortex, form a network involved in lexical representations (Lau et al., 2008) that has been associated with the N400. But then, as mentioned before, no concise time windows were used to determine these brain areas (Lau et al., 2008).

Relationships of musical to domain general cognitive mechanisms

Whereas our data revealed that working memory capacity predicted transgression discrimination in music, fluid intelligence did not. Working memory is crucial for the accurate performance of our task. In order to judge the appropriateness of the endings of the musical pieces, participants should keep the preceding context activated in memory and match it with the ending. Meanwhile they must compare the appropriateness of the ending to existing knowledge, implicitly or explicitly acquired. Interestingly this exactly corresponds to the integration function of the N400 described above and thus intimately links processing of incongruity of meaning in music and language. Therefore this finding illustrates that musical cognition shares certain basic features with domain general cognitive functions.

The fact that the Raven scores did not show a relationship with the ability to discriminate tonal violation suggests that on a higher level of cognition, musical processing and general fluid intelligence are independent. Nevertheless, there is a growing literature that shows that specifically language abilities (Besson et al., 2007) but also IQ (Moreno et al., 2011; Oechslin et al., 2013a; Schellenberg, 2004) and even mathematical skills (Schlaug et al., 2005) may be boosted by musical training. As in our sample, average level of musical education was modest; it is thus plausible that we did not find a correlation between fluid intelligence and tonal violation discrimination in music.

Conclusion

We observed a centro-posterior negativity (CPN) reminiscent of the semantic mismatch N400 in schoolchildren in response to refined syntactical incongruities in complex musical stimuli. Pre-motor and superior parietal areas supported the CPN that may play a role in incongruity discrimination and information integration. In a rich and varying musical context, integration of subtly incongruous musical elements may imply analysis of structure but also of meaning, at least in children.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.09.047>.

References

- Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., Besson, M., 2010. Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *J. Cogn. Neurosci.* 22, 2555–2569.
- Benau, E., Morris, J., Couperus, J.W., 2011. Semantic processing in children and adults: incongruity and the N400. *J. Psycholinguist. Res.* 40, 225–239.
- Besson, M., Faïta, 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., Schon, D., Moreno, S., Santos, A., Magne, C., 2007. Influence of musical expertise and musical training on pitch processing in music and language. *Restor. Neurol. Neurosci.* 25, 399–410.
- Besson, M., Frey, A., Aramaki, M., 2011. Is the distinction between intra- and extra-musical meaning implemented in the brain? Comment on “Towards a neural basis of processing musical semantics” by Stefan Koelsch. *Phys. Life Rev.* 8, 112–113 (discussion 125–118).
- Bigand, E., Poulin, B., Tillmann, B., Madurell, F., D’Adamo, D.A., 2003. Sensory versus cognitive components in harmonic priming. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 159–171.
- Brandt, A., Gebrian, M., Slevc, L.R., 2012. Music and early language acquisition. *Front. Psychol.* 3, 327.
- Britz, J., Landis, T., Michel, C.M., 2009. Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb. Cortex* 19, 55–65.
- Britz, J., Diaz Hernandez, L., Ro, T., Michel, C.M., 2014. EEG-microstate dependent emergence of perceptual awareness. *Front. Behav. Neurosci.* 8, 163.
- Brown, S., Martinez, M.J., Parsons, L.M., 2006. Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur. J. Neurosci.* 23, 2791–2803.
- Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011, 813870.
- Castle, P.C., Van Toller, S., Milligan, G.J., 2000. The effect of odour priming on cortical EEG and visual ERP responses. *Int. J. Psychophysiol.* 36, 123–131.
- Cox, A., 2011. Embodying music: principles of the mimetic hypothesis. *Music Theory Online* 17.
- Daltrozzo, J., Schon, D., 2009. Conceptual processing in music as revealed by N400 effects on words and musical targets. *J. Cogn. Neurosci.* 21, 1882–1892.
- Featherstone, C.R., Morrison, C.M., Waterman, M.G., MacGregor, L.J., 2013. Semantics, syntax or neither? A case for resolution in the interpretation of N500 and P600 responses to harmonic incongruities. *PLoS ONE* 8, e76600.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84.
- Friederici, A.D., Steinhauer, K., Frisch, S., 1999. Lexical integration: sequential effects of syntactic and semantic information. *Mem. Cognit.* 27, 438–453.
- Frishkoff, G.A., Tucker, D.M., Davey, C., Scherg, M., 2004. Frontal and posterior sources of event-related potentials in semantic comprehension. *Brain Res. Cogn. Brain Res.* 20, 329–354.
- Glenberg, A.M., 2010. Embodiment as a unifying perspective for psychology. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 586–596.
- Grave de Peralta Mendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S.L., 2004. Electrical neuroimaging based on biophysical constraints. *NeuroImage* 21, 527–539.
- Grech, R., Cassar, T., Muscat, J., Camilleri, K.P., Fabri, S.G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., Vanrumste, B., 2008. Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5, 25.
- Grégoire, J., 2009. L'examen clinique de l'intelligence de l'enfant: fondements et pratique du WISC-IV. Mardaga, Sprimont.
- Hamalainen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Herholz, S.C., Halpern, A.R., Zatorre, R.J., 2012. Neuronal correlates of perception, imagery, and memory for familiar tunes. *J. Cogn. Neurosci.* 24, 1382–1397.
- Isaacs, K.L., Barr, W.B., Nelson, P.K., Devinsky, O., 2006. Degree of handedness and cerebral dominance. *Neurology* 66, 1855–1858.
- James, C.E., 2012. Music and language processing share behavioral and cerebral features. *Front. Psychol.* 3, 52.
- James, C.E., Britz, J., Vuilleumier, P., Hauert, C.A., Michel, C.M., 2008. Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *NeuroImage* 42, 1597–1608.
- James, C.E., Dupuis-Lozeron, E., Hauert, C.-A., 2012. Appraisal of musical syntax violations by primary school children. *Swiss J. Psychol.* 71, 161–168.
- Jankélévitch, V., Abbate, C., 2003. Music and the Ineffable. Princeton University Press, Princeton, NJ.
- Jentschke, S., Koelsch, S., 2009. Musical training modulates the development of syntax processing in children. *NeuroImage* 47, 735–744.
- Knosche, T.R., Neuhaus, C., Hauelsen, J., Alter, K., Maess, B., Witte, O.W., Friederici, A.D., 2005. Perception of phrase structure in music. *Hum. Brain Mapp.* 24, 259–273.
- Koelsch, S., 2011. Toward a neural basis of music perception — a review and updated model. *Front. Psychol.* 2, 110.
- Koelsch, S., 2012. Brain and Music. Wiley-Blackwell, Chichester, West Sussex; Hoboken, NJ.
- Koelsch, S., Mulder, J., 2002. Electric brain responses to inappropriate harmonies during listening to expressive music. *Clin. Neurophysiol.* 113, 862–869.
- Koelsch, S., Gunter, T., Friederici, A.D., Schroger, E., 2000. Brain indices of music processing: “nonmusicians” are musical. *J. Cogn. Neurosci.* 12, 520–541.
- Koelsch, S., Gunter, T.C., Schroger, E., Tervaniemi, M., Sammler, D., Friederici, A.D., 2001. Differentiating ERAN and MMN: an ERP study. *Neuroreport* 12, 1385–1389.
- Koelsch, S., Grossmann, T., Gunter, T.C., Hahne, A., Schroger, E., Friederici, A.D., 2003. Children processing music: electric brain responses reveal musical competence and gender differences. *J. Cogn. Neurosci.* 15, 683–693.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., Friederici, A.D., 2004. Music, language and meaning: brain signatures of semantic processing. *Nat. Neurosci.* 7, 302–307.
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., Schlaug, G., 2005. Adults and children processing music: an fMRI study. *NeuroImage* 25, 1068–1076.
- Koelsch, S., Jentschke, S., Sammler, D., Mietschen, D., 2007. Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology* 44, 476–490.
- Koelsch, S., Kilches, S., Steinbeis, N., Schelinski, S., 2008. Effects of unexpected chords and of performer’s expression on brain responses and electrodermal activity. *PLoS ONE* 3, e2631.
- Koelsch, S., Rohrmeier, M., Torrecuso, R., Jentschke, S., 2013. Processing of hierarchical syntactic structure in music. *Proc. Natl. Acad. Sci. U. S. A.* 110, 15443–15448.
- Koenig, T., Stein, M., Grieder, M., Kottlow, M., 2014. A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain Topogr.* 27, 72–83.
- Kondakor, I., Lehmann, D., Michel, C.M., Brandeis, D., Kochi, K., Koenig, 1997. Prestimulus EEG microstates influence visual event-related potential microstates in field maps with 47 channels. *J. Neural Transm.* 104, 161–173.
- Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647.
- Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Lamont, A., Webb, R., 2010. Short- and long-term musical preferences: what makes a favourite piece of music? *Psychol. Music* 38, 222–241.
- Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48, 609–621.
- Lehmann, D., Ozaki, H., Pal, I., 1987. EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67, 271–288.
- Lehmann, D., Faber, P.L., Galderisi, S., Herrmann, W.M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R.D., Saito, N., Wackermann, J., Winterer, G., Koenig, T., 2005. EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: a multi-center study. *Psychiatry Res. Neuroimaging* 138, 141–156.
- Macmillan, N.A., Creelman, C.D., 1997. d’plus: a program to calculate accuracy and bias measures from detection and discrimination data. *Spat. Vis.* 11, 141–143.
- 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.
- Magne, C., Schon, D., Besson, M., 2006. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* 18, 199–211.
- Meyer, L.B., 1956. *Emotion and Meaning in Music*. The University of Chicago Press, Chicago.
- Meyer, G.F., Greenlee, M., Wuergler, S., 2011. Interactions between auditory and visual semantic stimulus classes: evidence for common processing networks for speech and body actions. *J. Cogn. Neurosci.* 23, 2291–2308.
- Michel, C.M., Brandeis, D., 2009. Data acquisition and pre-processing standards for electrical neuroimaging. In: Michel, C.M., Koenig, T., Brandeis, D., Gianotti, L.R.R., Wackermann, J. (Eds.), *Electrical Neuroimaging*. Cambridge University Press, Cambridge, pp. 79–92.
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004a. EEG source imaging. *Clin. Neurophysiol.* 115, 2195–2222.
- Michel, C.M., Seeck, M., Murray, M.M., 2004b. The speed of visual cognition. *Suppl. Clin. Neurophysiol.* 57, 617–627.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., Besson, M., 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* 19, 712–723.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Cepeda, N.J., Chau, T., 2011. Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433.

- Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. *J. Neurosci.* 26, 1293–1302.
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20, 249–264.
- Neuhaus, C., Knosche, T.R., Friederici, A.D., 2006. Effects of musical expertise and boundary markers on phrase perception in music. *J. Cogn. Neurosci.* 18, 472–493.
- Niedeggen, M., Rosler, F., Jost, K., 1999. Processing of incongruous mental calculation problems: evidence for an arithmetic N400 effect. *Psychophysiology* 36, 307–324.
- Nittono, H., Bito, T., Hayashi, M., Sakata, S., Hori, T., 2000. Event-related potentials elicited by wrong terminal notes: effects of temporal disruption. *Biol. Psychol.* 52, 1–16.
- Oechslin, M.S., Descloux, C., Croquelois, A., Chanal, J., Van De Ville, D., Lazeyras, F., James, C.E., 2013a. Hippocampal volume predicts fluid intelligence in musically trained people. *Hippocampus* 23, 552–558.
- Oechslin, M.S., Van De Ville, D., Lazeyras, F., Hauert, C.A., James, C.E., 2013b. Degree of musical expertise modulates higher order brain functioning. *Cereb. Cortex* 23, 2213–2224.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665.
- Patel, A.D., 2003. Language, music, syntax and the brain. *Nat. Neurosci.* 6, 674–681.
- Perrachione, T.K., Fedorenko, E.G., Vinke, L., Gibson, E., Dilley, L.C., 2013. Evidence for shared cognitive processing of pitch in music and language. *PLoS One* 8, e73372.
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187.
- Pinker, S., 1997. *How the Mind Works*. W. W. Norton & Company, New York.
- Platel, H., Price, C., Baron, J.C., Wise, R., Lambert, J., Frackowiak, R.S., Lechevalier, B., Eustache, F., 1997. The structural components of music perception. A functional anatomical study. *Brain* 120 (Pt 2), 229–243.
- Plomp, G., Michel, C.M., Herzog, M.H., 2010. Electrical source dynamics in three functional localizer paradigms. *NeuroImage* 53, 257–267.
- Plomp, G., Roinishvili, M., Chkonia, E., Kapanadze, G., Kereselidze, M., Brand, A., Herzog, M.H., 2013. Electrophysiological evidence for ventral stream deficits in schizophrenia patients. *Schizophr. Bull.* 39, 547–554.
- Polich, J., Criado, J.R., 2006. Neuropsychology and neuropharmacology of P3a and P3b. *Int. J. Psychophysiol.* 60, 172–185.
- Poulin-Charronnat, B., Bigand, E., Koelsch, S., 2006. Processing of musical syntax tonic versus subdominant: an event-related potential study. *J. Cogn. Neurosci.* 18, 1545–1554.
- Pulli, K., Karma, K., Norio, R., Sistonen, P., Goring, H.H., Jarvela, I., 2008. Genome-wide linkage scan for loci of musical aptitude in Finnish families: evidence for a major locus at 4q22. *J. Med. Genet.* 45, 451–456.
- Raven, J., Court, J., Raven, J., 1998. *Manuel des Raven, Section 3, Progressive Matrices Standard (PM 38)* (L. Marcou, Trans.). Editions et Applications Psychologiques, Paris.
- Razali, N.M., Wah, Y.B., 2011. Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. *J. Stat. Model. Anal.* 2, 21–33.
- 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.
- Rihs, T.A., Tomescu, M.I., Britz, J., Rochas, V., Custo, A., Schneider, M., Debbane, M., Eliez, S., Michel, C.M., 2013. Altered auditory processing in frontal and left temporal cortex in 22q11.2 deletion syndrome: a group at high genetic risk for schizophrenia. *Psychiatry Res.* 212, 141–149.
- Sammler, D., Novembre, G., Koelsch, S., Keller, P.E., 2013. Syntax in a pianist's hand: ERP signatures of “embodied” syntax processing in music. *Cortex* 49, 1325–1339.
- Satoh, M., Takeda, K., Nagata, K., Hatazawa, J., Kuzuhara, S., 2001. Activated brain regions in musicians during an ensemble: a PET study. *Brain Res. Cogn. Brain Res.* 12, 101–108.
- Schellenberg, E.G., 2004. Music lessons enhance IQ. *Psychol. Sci.* 15, 511–514.
- Schlaug, G., Norton, A., Overy, K., Winner, E., 2005. Effects of music training on the child's brain and cognitive development. *Ann. N. Y. Acad. Sci.* 1060, 219–230.
- Schon, D., Francois, C., 2011. Musical expertise and statistical learning of musical and linguistic structures. *Front. Psychol.* 2, 167.
- Schubotz, R.I., 2007. Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218.
- Spinelli, L., Andino, S.G., Lantz, G., Seeck, M., Michel, C.M., 2000. Electromagnetic inverse solutions in anatomically constrained spherical head models. *Brain Topogr.* 13, 115–125.
- Srebro, R., 1996. A bootstrap method to compare the shapes of two scalp fields. *Electroencephalogr. Clin. Neurophysiol.* 100, 25–32.
- Steinbeis, N., Koelsch, S., 2008. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb. Cortex* 18, 1169–1178.
- Steinbeis, N., Koelsch, S., Sloboda, J.A., 2006. The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *J. Cogn. Neurosci.* 18, 1380–1393.
- Swain, J., 1997. *Musical Languages*. W.W. Norton & Company, New York.
- Tibshirani, R., Walther, G., 2005. Cluster validation by prediction strength. *J. Comput. Graph. Stat.* 14, 511–528.
- Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D., von Cramon, D. Y., 2006. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *NeuroImage* 31, 1771–1782.
- Trainor, L.J., Trehub, S.E., 1994. Key membership and implied harmony in Western tonal music: developmental perspectives. *Percept. Psychophys.* 56, 125–132.
- Vaughan Jr., H.G., 1982. The neural origins of human event-related potentials. *Ann. N. Y. Acad. Sci.* 388, 125–138.
- Wechsler, D., 1974. *Manual for the Wechsler Intelligence Scale for Children—Revised*. The Psychological Corporation, New York.
- Wechsler, D., 2005. *WISC-IV: échelle d'intelligence de Wechsler pour enfants et adolescents*. ECPA, Paris.
- Ye, Z., Donamayor, N., Munte, T.F., 2014. Brain network of semantic integration in sentence reading: insights from independent component analysis and graph theoretical analysis. *Hum. Brain Mapp.* 35, 367–376.
- Zatorre, R.J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558.