The human brain processes hierarchical structures of meter and harmony differently: evidence from musicians and nonmusicians

Article

Accepted Version

Sun, L., Thompson, W. F., Liu, F. ORCID: https://orcid.org/0000-0002-7776-0222, Zhou, L. and Jiang, C. (2020) The human brain processes hierarchical structures of meter and harmony differently: evidence from musicians and nonmusicians. Psychophysiology, 57 (9). e13598. ISSN 0048-5772 doi: 10.1111/psyp.13598 Available at https://centaur.reading.ac.uk/90399/

To link to this article DOI: http://dx.doi.org/10.1111/psyp.13598

Publisher: Wiley-Blackwell

The human brain processes hierarchical structures of meter and harmony differently: Evidence from musicians and nonmusicians

Running head: Brain responses to musical structures

Lijun Sun[1], William Forde Thompson[2],

Fang Liu[3], Linshu Zhou[4], Cunmei Jiang[4,5]

[1] Department of Psychology, College of Education, Shanghai Normal University, Shanghai, China, 200234

[2] Department of Psychology, Macquarie University, Sydney, Australia, 2109

[3] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK, RG6 6AL

[4] Music College, Shanghai Normal University, Shanghai, China, 200234

[5] Institute of Psychology, Shanghai Normal University, Shanghai, China, 200234

Corresponding author:

Cunmei Jiang

Music College, Shanghai Normal University, 100 Guilin Road, Xuhui District, Shanghai, China 200234

cunmeijiang@126.com

Abstract

Research into how the brain processes temporal structure has gained increasing attention, yet there is remarkably little understanding of how temporal and non-temporal structures are processed simultaneously. Using event-related potentials (ERPs), we examined how the brain responds to temporal (metric) and non-temporal (harmonic) structures in music simultaneously, and whether these processes are impacted by musical expertise. Fifteen musicians and 15 nonmusicians rated the degree of completeness of musical sequences with or without violations in metric or harmonic structures. In the single violation conditions, the ERP results showed that both musicians and nonmusicians exhibited an early right anterior negativity (ERAN) as well as an N5 to temporal violations ("when"), and only an N5-like response to non-temporal violations ("what") , which were consistent with the behavioral results. In the double violation condition, however, only the ERP results, but not the behavioral results, revealed a significant interaction between temporal and non-temporal violations at a later integrative stage, as manifested by an enlarged N5 effect compared to the single violation conditions. These findings provide the first evidence that the human brain uses different neural mechanisms in processing metric and harmonic structures in music, which may shed light on how the brain generates predictions for "what" and "when" events in the natural environment.

Keywords: *what* and *when* information, harmonic structure, metric structure, ERAN, N5, musical expertise

1 Introduction

Research into how the brain processes temporal structure has gained increasing attention (Dehaene, Meyniel, Wacongne, Wang, & Pallier, 2015), yet there is remarkably little understanding of how temporal and non-temporal structures are processed simultaneously, especially when such structures are organized in nested ways. Dynamic attending theory, proposed by Jones (Jones, 1976; Jones & Boltz, 1989), posits that attention becomes entrained to temporal events and this entrainment facilitates the processing of non-temporal events presented in phase with the temporal structure of auditory sequences. Although the theory has received some empirical support (e.g., Boltz, 1993; Jones, Johnston, & Puente, 2006; Jones, Moynihan, Mackenzie, & Puente, 2002; Prince, Schmuckler, & Thompson, 2009), evidence is lacking on how temporal and non-temporal events interact with each other in hierarchical sequences.

Given its emphasis on hierarchically structured metric (*when*) and harmonic (*what*) structure, music provides a unique window into how the brain simultaneously processes temporal and non-temporal structures (Cuddy, Cohen, & Mewhort, 1981; Fitch, 2013; Koelsch, 2013; Prince, Thompson, & Schmuckler, 2009; Russo, Thompson, & Cuddy, 2015; Simon, 1972). In Western tonal music, metric structures are hierarchically organized based on strong and weak beats (Jones, 2009; Patel, 2008; Prince, Thompson, et al., 2009), while harmonic structures are organized based on the stability of notes or chords (Krumhansl, 1990). When harmonic and metric hierarchies are misaligned, they can be difficult to process (Jones, 1987; Jones & Boltz, 1989; Prince, Thompson, et al., 2009). However, little else is known about how these two structures are processed simultaneously.

Behavioral studies have demonstrated that the processing of hierarchical harmonic structures is enhanced by regular, non-hierarchical temporal structures such as expected isochronous temporal events (Tillmann & Lebrun-Guillaud, 2006) or symmetric temporal structure (Bigand, Madurell, Tillmann, & Pineau, 1999; Boltz, 1989). Nevertheless, whether hierarchical temporal (metric) structures also facilitate the processing of hierarchical harmonic structures remains to be demonstrated. In one investigation, it was found that judgments of metric position were biased by tonal stability, but judgments of tonality were unaffected by metric position (Prince, Thompson, et al., 2009). However, this asymmetric influence could not reveal the exact relationship between metric and harmonic structural processing.

In the present investigation, we used ERPs to examine how metric and harmonic structures in the same musical sequence are processed, with the aim of elucidating how the brain simultaneously processes *when* and *what* information in hierarchically organized sequences. To our knowledge, there have been no electrophysiological studies examining how metric and harmonic structures are *simultaneously* processed in the brain. Rather, previous research has focused either on harmonic syntactic structures (e.g., Koelsch & Jentschke, 2010; Koelsch, Jentschke, Sammler, & Mietchen, 2007; Steinbeis & Koelsch, 2008; Zhou, Liu, Jiang, Jiang, & Jiang, 2019) or on rhythmic syntactic structures (Sun, Liu, Zhou, & Jiang, 2018). When brain responses are examined for harmonic structure, it has been reported that out-of-key chords or notes elicit an ERAN and an N5 (indexing integration of expectancy or the processing of intra-musical meaning) in both musicians and nonmusicians (e.g., Koelsch & Jentschke, 2010; Koelsch et al., 2007; Steinbeis & Koelsch, 2008). However, when in-key chords were used in the harmonic irregular condition, no ERAN was observed (Poulin-Charronnat, Bigand, & Koelsch, 2006), which suggests that the ERAN effect is modulated by the perceived psychological distance between the regular and irregular chords. Similarly, an N5-like effect has been observed when the scalp distribution of the N5 is modulated by task, i.e., shifting from the frontal distribution under an attended condition to the whole-scalp distribution under an unattended condition (Loui, Grent, Torpey, & Woldorff, 2005). When brain responses are examined for rhythmic structure, it has been reported that rhythmic syntactic violations elicit an early right anterior negativity (ERAN, indexing the processing of a musical expectancy violation) in musicians, but not in nonmusicians (Sun et al., 2018). Unlike those studies which examined either metric or harmonic regularity, our study manipulated both metric and harmonic regularities simultaneously at the end of musical sequences. Thus, the first aim of our study was to determine how the brain processes metric and harmonic structures simultaneously.

The second aim of our study was to examine whether *musical expertise* might benefit the simultaneous processing of metric and harmonic structures in music. Although listeners who grow up in the Western tonal music environment should possess the ability to process metric and harmonic structures regardless of their musical training background (Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch et al., 2007), previous studies have shown that musicians outperform nonmusicians in processing harmonic (Jentschke & Koelsch, 2009; Koelsch & Jentschke, 2008; Koelsch, Schmidt, & Kansok, 2002) and rhythmic syntactic structures (Sun et al., 2018), when these structures are examined separately. Exploring the role of learning in the simultaneous processing of metric and harmonic structures in music would contribute to our knowledge about the effects of musical training on musical structure processing.

Given the asymmetric interaction between tonal stability and metric position on the stability ratings of musical sequences (Prince, Thompson, et al., 2009), we expected that metric and harmonic violations would elicit different neural responses in the brain, and musicianship might also influence such responses.

2 Materials and methods

2.1 Participants

Fifteen musicians (Mage = 22.47 years, SD = 2.39, 11 females and 4 males) and 15 nonmusicians (Mage = 24.07 years, SD = 2.19, 11 females and 4 males) participated in the experiment. In order to determine whether our sample size was sufficient to detect a between-group effect, we estimated the power of our analysis to detect an effect using the data from our previous study (Sun et al., 2018). Two simulated datasets were created by randomly choosing 15 participant datasets (with replacement) separately from the musician and nonmusician groups in the previous study. The group ERP differences were then calculated using cluster-based random permutation tests (Maris & Oostenveld, 2007). After repeating this procedure 2000 times, power was estimated by calculating the proportion of repetitions yielding significant results to all repetitions. The result revealed an estimated power of 0.82 with 15 participants in each group, confirming that our sample size was sufficient to detect a between-group effect in the present study.

All participants in our experiment were native Chinese college students. The musicians had high proficiency in Western tonal music and had received on average 14 years of formal instrumental training (range: 9 to 18) on piano, violin, viola, accordion, cello or erhu (a traditional Chinese instrument), practicing on average more than 3 hours per day. The nonmusicians had never received extracurricular music training, although the compulsory education curriculum in China contains music lessons of 40-45 minutes per week. However, they were typically familiar with Western tonal music since listening to tonal music is an essential experience for Mandarin listeners (Jiang, Liu, & Wong, 2017; Wong, Roy, & Margulis, 2009). All participants were right-handed and reported no history of neurological, major medical or psychiatric disorders or hearing impairments. Ethical approval was obtained from Shanghai Normal University, and all participants signed a written consent form before the experiment was conducted.

2.2 Experimental design and statistical analysis

Rationale for the experimental design. Using a similar design in our previous study (Sun et al., 2018), we created irregular sequences with a syncopated ending chord, in which sound pressure level (SPL) accented chord was placed at a metrically weak rather than strong position. To reduce demands on working memory, chorale sequences were five-bars rather than eight-bars as in our previous study (Sun et al., 2018).

Because our goal was to evaluate and compare brain responses to harmonic and metric structure when they are combined in a musical sequence, we designed our stimuli such that manipulations of harmonic and metric structure were equated for discriminability. To this end, we conducted Pretest 1 to establish roughly equal discriminability between harmonic and metric structures. As our aim was to understand how the brain extracts and processes metric and harmonic regularities, we did not provide participants with explicit metrical cues. Therefore, we conducted Pretest 2 to determine whether the stimuli were perceived in 2/4, 3/8, or additive meter in order to ensure stimulus validity.

Pretests In Pretest 1, we asked 11 nonmusicians to rate the similarity between different versions of the same original sequence on a 7-point Likert scale, with regular/irregular metric and/or harmonic structures in a pair. For metric structures, regular metric sequences ended with a chord at the first metrically strong position, and irregular sequences ended with a syncopated chord, in which an accented chord was placed at a metrically weak rather than strong position. For harmonic structures, regular harmonic sequences ended with tonic chords, and irregular harmonic sequences ended with either supertonic chords or double dominant chords. Similarity ratings indicated that discriminability of regular and irregular metric endings closely matched the discriminability of tonic (regular) and supertonic (irregular) harmonic conditions (see Supplemental files). Therefore, the supertonic chords were chosen over double dominant chords as the harmonically irregular endings for the formal experiment (see Supplemental files, Pretest 1).

In Pretest 2, we tested 11 nonmusicians and 11 musicians who did not take part in the previous pretest or the formal experiment, with a 3 meter (2/4, 3/8, additive) × 2 group (musicians, nonmusicians) two-factor mixed design. Ten four-part chorale sequences were chosen as potential stimuli based on the results from Pretest 1. We then manipulated the regularities of the metric (rhythmic accents at metrically strong vs. weak positions) and harmonic (tonic vs. supertonic chords) events at the end of these sequences, which served as standard sequences. The comparison sequences were melodies derived from the top voices of the standard sequences and were presented in 2/4-, 3/8-, and additive meters, respectively. In order to provide a cue of meter to the participants, the first beats of each measure in the comparison sequences were intentionally accented. Each standard sequence was paired with three different comparison sequences, in 2/4-, 3/8-, or additive meter. Participants were required to make a same/different judgment of the paired sequences based on metrical similarity.

The percentages of ‘yes’ response to the 2/4-, 3/8-, and additive meters were calculated for each group. Repeated measures analyses of variance (ANOVAs) with meter (2/4, 3/8, additive) as the within-subjects factor and group (musicians, nonmusicians) as the between-subjects factor were conducted. The results showed that the main effect of meter was significant (p = .004). Neither the main effect of group (p = .99) nor the interaction between meter and group (p = .65) was significant. Post hoc comparative analysis showed that participants perceived the sequences more frequently in 2/4 meter than in 3/8 and additive meters (2/4: M = 60.23, 3/8: M = 41.36, additive: M = 46.48). Single sample t-tests showed that the perception of 2/4-meter, but not that of 3/8 or additive meter, was significantly higher than the chance level (p = .037). These results confirmed that our stimuli were perceived in 2/4 meter, with no difference between musicians and nonmusicians (see Supplemental files, Pretest 2).

Stimulus. Based on the results of the two pretests, we established equivalence in hierarchical regularities of the metric (rhythmic accents at metrically strong vs. weak positions) and harmonic (tonic vs. supertonic chords) events at the end of the sequences. An example of the experimental stimuli in the formal experiment are shown in Figure 1. We included four experimental conditions to examine regular and irregular metric and harmonic structures: 1) regular metric and harmonic structures (regular); 2) regular metric and irregular harmonic structures (single violation); 3) irregular metric and regular harmonic structures (single violation); and 4) irregular metric and harmonic structures (double violation). A regular metric sequence is expected to end with a rhythmic accent at a metrically strong position, while a regular harmonic sequence is expected to be terminated by an authentic cadence (V-I in major keys). As in previous studies, we focused on the ERAN and N5, the two main ERP components associated with the processing of musical structures.

Insert Figure 1, about here.

There were 5 five-bar original chorale sequences in 2/4 meter in the formal experiment. We controlled for the frequency of occurrence of the critical ending events for both harmonic and metric structures so as to avoid the effect of sensory novelty on musical structural processing (Koelsch et al., 2007). Because of the relatively high number of irregular ending chords (supertonic chords) in our sequences, the harmonic progressions usually developed around the category of subdominant function. However, on the whole, the harmonic progressions of our musical sequences conformed with tonal conventions of Western music. Taking the sequence in Figure 1 as an example, it started with tonic-dominant-tonic harmonic progression, developed around the category of subdominant function, and finally ended with the authentic cadence.

Each sequence had four versions, with regular/irregular metric and/or harmonic structure, thus forming 20 sequences. These sequences were transposed to 12 major keys, thus resulting in 240 sequences. All stimuli were created at a tempo of 82 quarter notes per minute using the Sibelius 7.5 software (Avid Tech. Inc.), and exported with Steinway piano timbre by Cubase 5.1's inbuilt Kontact 5.4. The loudness of each sequence was normalized to -3 dB using Adobe Audition 3.0.

Procedure. 240 sequences were presented in a pseudorandom order with the constraints that a given type of ending was not repeated more than three times in succession, and that consecutive sequences were neither from the same original sequence nor in the same key. All stimuli were presented through loudspeakers (Edifier, R101V).

Participants were instructed to first rely on intuition to judge whether the musical events in a sequence followed one another in an expected manner, and then to rate the degree of completeness of each sequence on a 5-point Likert scale, with 1 indicating the lowest degree of completeness and 5 the highest degree of completeness. They were encouraged to use the full range of the response scale. The completeness judgments were not solely based on the ends of musical sequences. Instead, such judgements required global integration of the whole musical sequence (Tillmann & Lebrun-Guillaud, 2006). Therefore, participants were instructed to compare musical expectation with viewers' expectation for movie plots, given that expectation for future sounds is tailored to a particular context (Huron, 2006). That is, when watching a movie, one often builds up expectations from one plot to the next. Expected plots in a movie would lead to the resolution of these plots and make viewers feel fulfilled and relaxed. If a movie ended with an unexpected plot, the viewers would feel a sense of incompleteness, and vice versa. Following the instruction, eight practice trials were given to familiarize the participants with the stimuli and procedure. Before the formal experiment, participants were also given the chance to adjust the volume of the stimuli to their comfortable listening level.

ERP recording and analysis. EEG data were recorded using a Neuroscan Synamps amplifier with 64 standard scalp locations according to the International 10–20 system. To measure eye movements and eye-blinks, electrodes were placed above and below the left eye as well as the outer canthi of both eyes. Signals were digitized with a sampling rate of 500 Hz, and filtered using a 0.05 Hz low cutoff and a

100 Hz high cutoff. The left mastoid electrode served as on-line reference. After the measurements, the data were referenced offline to the algebraical mean of left and right mastoid electrodes, with a 0.1 to 30 Hz band-pass filter (24-dB/oct slope) using NeuroScan software 4.5. We realigned and time-locked the recording time window to the last chords. Epochs of 1200 msec including a 200 msec pre-stimulus baseline period were averaged. Trials were rejected from the data when artifacts exceeded the amplitude of ±75 μV in any channel. On average, 86% of the trials were kept, similarly across all four conditions (all *p*s > .05).

As indicated by the rectangles in Figure 1, the ERPs were extracted from the first chord of the last bar (5.85 s after the music onset) for metrically regular endings and from the last chord of the penultimate bar (5.49 s after the music onset) for metrically irregular endings. These sequence lengths were sufficient for participants to establish the regularity representations of harmony and meter in our experiment, given that the minimum length required was a minimum of about 1 s (Andreou, Griffiths, & Chait, 2015; Lebrun-Guillaud & Tillmann, 2007).

We performed cluster-based random permutation tests (Maris & Oostenveld, 2007) using the MATLAB toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011), in order to control for Type I error rates due to multiple comparisons across different conditions in ERP studies (Luck & Gaspelin, 2017; Wang, Zhu, Bastiaansen, Hagoort, & Yang, 2013). All spatially adjacent data samples exceeding a preset significance level (5% here) were grouped into clusters. For each cluster, the sum of the *t* statistics was used as the cluster-level test statistic. Then a null distribution was created with the assumption of no difference between conditions. This distribution was obtained by randomly assigning the conditions in participants 1000 times and then the largest cluster level statistic was calculated for each randomization. Finally, the observed cluster-level test statistics were compared against the null distribution, and clusters falling in the highest or lowest 2.5th percentile were considered significant.

We examined the early right anterior negativity (ERAN) and the N5 as the ERP components for the processing of musical structure (e.g., Koelsch et al., 2000; Koelsch & Jentschke, 2010). Based on previous findings (Koelsch et al., 2000; Loui et al., 2005; Poulin-Charronnat et al., 2006) and visual inspection, we selected two time windows for analysis: (1) 100–250 msec after the chord onset to test for ERAN effects, and (2) 300–800 msec after the chord onset to test for N5 effects. The mean amplitudes of the trials in each condition were computed by including all but the reference electrodes.

In order to examine whether musicianship benefits the processing of metric and harmonic structures, we first computed the ERP effects (irregular minus regular) in single and double violation conditions in each group separately. We then subtracted the ERP effects in the double violation condition from the single violation condition. Following the comparisons between these effects for each group, we examined the three-way interaction among group, harmonic structure, and metric structure using a cluster-based random permutation test.

In cases where there was no significant group effect, we pooled the data and then tested the two-way interaction between metric and harmonic structures, by comparing the ERP effects (irregular minus regular) between single and double violation conditions separately for the two syntactic structures. In the event that there was a significant two-way interaction, planned contrasts were then conducted between irregular and regular endings in the single and double violation conditions separately for metric and harmonic structures. When no significant two-way interaction was found, we then combined the data across the two conditions and tested the main effect of metric and harmonic structures, respectively.

3 Results

3.1 Behavioral results

Figure 2 displays the mean completeness ratings of the sequences. A three-way ANOVA was conducted with group (musicians, nonmusicians) as a between-subjects factor and metric (regular, irregular) and harmonic structure (regular, irregular) as within-subjects factors. There was a significant main effect of metric structure ($F_{(1, 28)} = 35.32$, $p < .001$, partial $\eta^2 = .56$, ANOVA), indicating that sequences with irregular metric structure were perceived as less complete than those with regular metric structure (regular: $M = 3.13$, $SD = .05$, irregular: $M = 2.75$, $SD = .06$). There was also a significant main effect of harmonic structure ($F_{(1, 28)} = 465.11$, $p < .001$, partial $\eta^2 = .94$, ANOVA), indicating that sequences with irregular harmonic structure were perceived as less complete than those with regular harmonic structure (regular: $M = 4.10$, $SD = .06$; irregular: $M = 1.78$, $SD = .08$). No other significant main effects or interactions were observed (*p*s > .06, ANOVA). These results indicated that both musicians and nonmusicians were able to perceive the difference in completeness between regular and irregular conditions for both metric and harmonic structures.

Insert Figure 2, about here.

3.2 EEG results

The ERAN at the 100 to 250 msec time window. A three-way interaction test among group, metric structure, and harmonic structure showed that there was no significant difference between groups (*p* = .881, cluster-based random permutation test). We then combined the data across the two groups and tested the two-way interaction between metric and harmonic structures by comparing the metric and harmonic effects in the single violation condition with those in the double violation condition. The results showed no interaction between metric and harmonic structure (*p* = .368, cluster-based random permutation tests).

In order to examine whether there were main effects of metric and harmonic structure, we performed two sets of analyses contrasting irregular and regular conditions for metric and harmonic structures. For harmonic structure, no significant cluster was observed (*p* = .166, cluster-based random permutation tests), suggesting that there was no significant difference between the two conditions. For metric structure, however, the test revealed a significant difference between regular and irregular endings from 145 to 229 msec (*p* = .004, cluster-based random permutation tests). This difference was most pronounced over frontal sensors. These findings indicate that violations in metric structure elicited an ERAN over the frontal area, as shown in Figure 3.

Insert Figure 3, about here.

The N5 at the 300-800 msec time window. A three-way interaction test among group, metric structure, and harmonic structure showed that there was no significant difference between groups (*p* = .238, cluster-based random permutation tests). Thus, the data were combined across groups and a two-way interaction test between metric and harmonic structures was conducted. As shown in Figure 4, the test revealed a significant difference between single and double violation conditions from 515 to 800 msec (*p* = .004, cluster-based random permutation tests), which was most pronounced over right-dominant parietal-occipital sensors, indicating an interaction between the integration of metric and harmonic structures. A larger negativity was elicited in the double than the single violation condition for both metric and harmonic structures, suggesting that harmonic structural integration was affected by the simultaneous presence of a metric irregularity, and vice versa.

Insert Figure 4, about here.

As shown in Figure 5, planned contrasts between irregular and regular harmonic endings in the regular metric condition showed a significant difference from 300 to 647 msec (*p* = .002, cluster-based random permutation tests), which was most pronounced over left parietal-occipital sensors. Planned contrasts between irregular and regular harmonic endings in the irregular metric condition showed a significant difference from 335 to 774 msec (*p* = .002, cluster-based random permutation tests), which was most pronounced over parietal-occipital sensors.

Insert Figure 5, about here.

As shown in Figure 6, planned contrasts between irregular and regular metric endings in the regular harmonic condition showed a significant difference from 300 to 800 msec (*p* = .002, cluster-based random permutation tests), which was most pronounced over frontal sensors. Planned contrasts between irregular and regular metric endings in the irregular harmonic condition showed a significant difference from 347 to 800 msec (*p* = .002, cluster-based random permutation tests), which was present across the whole-scalp sensors.

Insert Figure 6, about here.

----

4 Discussion

Using ERPs, we examined how the brain simultaneously processes metric and harmonic structures in music, and whether these processes are impacted by musical expertise. We manipulated both metric and harmonic regularities simultaneously at the end of musical sequences, given that the hierarchies in music are typically represented in cadence. The results revealed a significant difference in neural responses to metric and harmonic structures in musical sequences that contain both forms. In particular, even when the discriminability of violations in metric and harmonic structures was equated, an ERAN was observed in response to metric but not harmonic violations. Given that the ERAN reflects automatic cognitive processing of musical structures (Koelsch et al., 2000; Koelsch & Jentschke, 2008; Koelsch et al., 2002), this finding suggests automatic processing of metric but not harmonic structure at an early processing stage. However, at the later integrative stage, we observed an interaction between metric and harmonic structures, as evidenced by an enlarged N5 effect in the double violation condition. Planned comparisons revealed that violations of metric structure elicited an N5 and violations of harmonic structure elicited an N5-like response in both single and double violation conditions. Furthermore, nonmusicians showed comparable neural responses as musicians to concurrent metric and harmonic structural violations.

Our first main finding is that an ERAN was observed in response to metric but not to harmonic structural violations. This ERAN effect was evoked by a musical structural violation and hence reflects musical structure building (Koelsch, 2013). It cannot be attributed to sensory novelty as we carefully controlled for the frequency of occurrence of the final events in our stimuli. That is, sequences ending with syncopated and nonsyncopated chords in our study were presented with equal probability, and the frequencies of occurrence of the regular and irregular metric chords were the same in each sequence. Thus, this ERAN reflects automatic cognitive processing of musical metric structure (Sun et al., 2018), which is consistent with previous findings that musical structural processing operates in the absence of attention (Koelsch et al., 2000; Koelsch & Jentschke, 2008; Koelsch et al., 2002), although it can also be modulated by demands of attention (Loui et al., 2005; Maidhof & Koelsch, 2011).

The differential results in the ERAN response between metric and harmonic structures may be due to the predictive feature of temporal processing, which might have contributed to the ERAN effect in metric structure. Indeed, beat perception is essentially a predictive rather than a reactive process. The brain is capable of making highly accurate temporal predictions about the timing of upcoming beats (Rankin, Large, & Fink, 2009; Repp, 2005; Repp & Su, 2013), even when the movement is absent (see Patel & Iversen, 2014). Such predictive processing is modulated by dynamic attention allocation (Jones & Boltz, 1989), which is correlated with neural oscillations (Large & Snyder, 2009). Indeed, neural oscillations reach their maximum prior to the occurrence of the next beat, which underpins prediction accuracy during timing processing (Arnal, Doelling, & Poeppel, 2014; Fujioka, Ross, & Trainor, 2015; Fujioka, Trainor, Large, & Ross, 2012).

The absence of an ERAN response to harmonic violations in our study is consistent with results reported by Poulin-Charronnat et al. (2006), where neither musicians nor nonmusicians showed an ERAN to harmonic violations using in-key chords (subdominant chords). However, other studies have reported an ERAN effect elicited by irregular harmonic endings (e.g., Koelsch et al., 2000; Koelsch & Jentschke, 2008; Koelsch et al., 2007; Koelsch & Mulder, 2002; Patel, 1998). This discrepancy may be explained by the difference in stimuli used across different studies. Firstly, the harmonic ERAN seems to depend on the saliency of the ending chords in the irregular condition. In our study, we used in-key (subdominant chords) rather than out-of-key chords as the irregular endings, as in the study by (Poulin-Charronnat et al., 2006). Most likely, an ERAN was observed for out-of-key chords because they are far more irregular and acoustically divergent than in-key chords (e.g., Koelsch et al., 2000; Koelsch & Jentschke, 2008; Koelsch et al., 2007; Koelsch & Mulder, 2002; Patel, 1998). Secondly, rhythmic/harmonic complexity may also modulate the harmonic ERAN. Unlike previous studies, we used sequences with varied rhythm rather than isochronous chords. High rhythmic complexity is thought to delay the latency of the harmonic ERAN (Koelsch & Jentschke, 2008; Koelsch & Mulder, 2002; Patel, 1998), but in our study may have had an even more disruptive impact on this neural response, making it harder for the brain to detect. Similarly, chorale sequences in our study contained more chords than those used in previous studies (e.g., Koelsch et al., 2000; Koelsch & Jentschke, 2008), which made our task more complex than previous ones. One or more of these factors may account for the absence of the harmonic ERAN in our study.

Our second main finding is that metric and harmonic structures were processed in an interactive manner at the later integrative stage. When only metric violations were present, our participants exhibited a right lateralized frontal N5 effect, an index of the integration of musical structure. This finding is consistent with previous findings that the N5 effect reflects harmonic structural integration, which is modulated by the degree of fit with regard to the previous musical context (e.g., Koelsch & Jentschke, 2010; Koelsch et al., 2007; Steinbeis & Koelsch, 2008). The N5 component signifies the difference in expectedness between regular and irregular cadences. Therefore, given that our experimental manipulation was based on musical expectations, the ERP effects we observed reflected the difference in expectedness between musical sequences.

In response to harmonic violations alone, our participants exhibited a posterior N5-like effect. Although this N5-like effect was induced in a similar time window as the metric N5, it showed a central-parietal, rather than a right lateralized frontal distribution. Thus, we referred to it as an N5-like effect, rather than an N5. Such a posterior effect may be due to the increased difficulty in structural integration. Specifically, previous studies have demonstrated a frontal N5 effect using real music containing varied rhythm and tonic versus out-of-key chords in the regular and irregular conditions (Koelsch & Mulder, 2002) or using isochronous chord sequences ending with tonic or in-key chords (Koelsch & Jentschke, 2010). In the present study, musical sequences were not only with varied rhythm, but also ended with tonic or in-key (supertonic chords) chords, which increased the task difficulty. Because the perceived psychological distance between out-of-key and tonic chords is greater than that between in-key (supertonic chords) and tonic chords, the former would be easier to perceive than the latter. This increased difficulty might have resulted in the posterior N5-like effect in the present study. On the other hand, compared with the metric N5 effect, this harmonic N5-like effect might be due to the absence of the harmonic ERAN at the early processing stage. That is, early automatic structural processing may enhance listeners' sensitivity to structural integration of metric structures at a later stage. However, this possibility will need to be validated in future studies.

In the double violation condition, concurrent metric and harmonic processing modulated the N5 and N5-like effects, as the N5 effect elicited by metric violations was enlarged by irregular harmonic endings and the N5-like effect elicited by harmonic violations was enlarged by irregular metric endings. As suggested by previous studies (Hagoort & Brown, 2000; Koelsch et al., 2000; Kutas & Kluender, 1994), this result may be attributed to the increased difficulty in integrating metric (or harmonic) structures in the presence of irregular harmonic (or metric) structures. This enlarged N5 suggests that more cognitive resources are required for the integration of double violations compared to single violations. Previous studies have demonstrated that regular non-hierarchical temporal structures facilitate pitch discrimination (Boltz, 1993; Jones et al., 2006; e.g., Jones et al., 2002; Prince, Schmuckler, et al., 2009) and the processing of hierarchical harmonic structures (Bigand et al., 1999; Boltz, 1989; Tillmann & Lebrun-Guillaud, 2006). Taken together, our findings extend previous results to suggest that regular metric structures that are hierarchically organized can also facilitate the processing of harmonic structures.

Although the irregular harmonic or metric endings enhanced the metric N5 or harmonic N5-like effect, these enhanced effects were not simply due to the sum of the metric and harmonic effects. An ANOVA with the N5 effect (double violation, sum), hemisphere (left, right), anteriority (anterior, posterior) as the within-subjects factors showed that the N5 effect elicited in the double violation condition was significantly larger than the sum of the metric and harmonic effects ($F_{(1,28)} = 8.07$, $p = .008$, partial $\eta^2 = .22$) (see Figure S1, Supplementary files). Given that the N5 is also sensitive to attentional resource allocation (Loui et al., 2005), the enhanced N5 and N5-like effects indicate that additional attentional resources may be required when integrating the irregular harmonic or metric ending into the previous musical context.

Our third main finding is that musical expertise is not associated with enhanced processing of combined metric and harmonic structures at the behavioral or neural levels. The absence of musical training effect cannot be attributed to limitations in the stimuli or sample size. For the metric structures, although we did not provide participants with explicit metrical cues, the results of Pretest 2 confirmed that our stimuli were perceived in 2/4 meter, rather than in 3/8, or additive meters. Furthermore, the absence of musical training effect in our study was unlikely due to our sample size, as our power estimation revealed that a sample of 15 musicians and 15 nonmusicians was sufficient to detect a between-group effect (see Methods 2.1).

Two possible reasons may account for the absence of such a musical training effect. First, for processing harmony, research has shown that nonmusicians can become remarkably sensitive to the regularities in tonal music through passive music exposure and perceptual learning (Jiang et al., 2017; Koelsch et al., 2000; Wong et al., 2009), which allows them to process musical structures (Bigand & Poulin-Charronnat, 2006; Zhou et al., 2019). Likewise, for processing meter, it has been suggested that a bias towards duple meter could result from listeners' daily exposure to Western tonal music (Trainor & Hannon, 2013), which is present even in 9-month-old infants (Bergeson & Trehub, 2006). Indeed, consistent with our findings, previous studies have also reported similarities in the way musicians and nonmusicians integrate harmonic structure (Jentschke & Koelsch, 2009; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Koelsch et al., 2002; Miranda & Ullman, 2007) and metric structure (Sun et al., 2018). These results corroborate our finding that nonmusicians and musicians exhibit a similar capacity to process musical structure.

Second, the absence of musical training effect may be related to task difficulty. Indeed, task difficulty modulates the effects of musical training on rhythmic production (Chen, Penhune, & Zatorre, 2008), beat perception (Bouwer, Werner, Knetemann, & Honing, 2016), and pitch contour processing (Schön, Magne, & Besson, 2004). Compared with our previous study (Sun et al., 2018), the present study used shorter musical sequences (five bars). Such a manipulation may make the task of metric structural processing easier, which may in turn result in the absence of the effect of musical expertise. Indeed, the difference between musicians and nonmusicians disappeared when only four-bar rhythmic sequences were presented (unpublished data). However, future research is needed to examine whether task difficulty and musical knowledge modulate the effect of musical training on the processing of musical structure.

In addition, the behavioral results are not completely consistent with the ERP results in the present study. This difference suggests that ERP and behavioral responses are sensitive to different cognitive operations (e.g., Kounios & Holcomb, 1992; Tillman & Wiens; Zhang, Guo, Ding, & Wang, 2006). Specifically, the ERP amplitudes such as the ERAN and N5 amplitudes to music proved to be sensitive to the detection of music structural irregularity and integration, respectively, whereas the behavioral responses may be more sensitive to the demands in judging the completeness of musical sequences. In other words, the ERAN and N5 in our study only reflected the operation of the mechanisms that manifest the properties of musical structures (i.e., detection and integration), whereas the behavioral responses might be sensitive to participants' decision-making processes and task-dependent strategies. On the other hand, the recording of the behavioral responses in the present study was outside of the ERP measurement windows used. Given that the ERPs were time-locked to the onsets of the final chords, the ERAN and N5 were measured from 100 to 250 msec and 300 to 800 msec after the final chord onsets. However, participants were required to rate the completeness after *the termination of the whole musical sequences*, rather than after *the onsets of the final chords*. Therefore, our behavioral results reflected the offline judgment of music-completeness, whereas our ERP results reflected the online processing of musical structure. This might have led to the difference in the interaction between metric and harmonic structures between these two types of results in our study.

5 Conclusion

This investigation provides the first electrophysiological evidence that neural responses to metric (when) and harmonic (what) structures are different in sequences that contain both forms of structure, even when changes are equated for discriminability. Our findings indicated automatic early processing of metric but not harmonic structure, and interactive processing of the two forms of structure at the later integration stage. These findings extend dynamic attending theory (Jones, 1976; Jones & Boltz, 1989) by indicating that regular, hierarchically organized temporal structures can facilitate the processing of hierarchically organized non-temporal structures, and vice versa.

It is well established that, in order to interact with and navigate in a constantly changing environment, the human brain must continuously make rapid and accurate predictions about *what* will happen next and *when* it will occur in temporal sequences. Our findings suggest that the human brain can automatically respond to the unpredicted *when* information (but not to the *what* information) at the early stage. At the later stage, however, the processing of the unpredicted *when* and *what* information seems to be interactive. Such neural responses are not affected by learning experience in individuals. Thus, our findings may shed light on how the brain generates predictions for “what” and “when” events in the natural environment.

References

Andreou, L.-V., Griffiths, T. D., & Chait, M. (2015). Sensitivity to the temporal structure of rapid sound sequences-An MEG study. *Neuroimage, 110*, 194-204. doi:https://www.doi.org/10.1016/j.neuroimage.2015.01.052

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2014). Delta-Beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex, 25*(9), 3077-3085. doi:https://www.doi.org/10.1093/cercor/bhu103

Bergeson, T. R., & Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception, 23*, 345-360. doi:https://www.doi.org/10.1525/mp.2006.23.4.345

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 184-197. doi:https://www.doi.org/10.1037/0096-1523.25.1.184

Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*, 100-130. doi:https://www.doi.org/10.1016/j.cognition.2005.11.007

Boltz, M. G. (1989). Rhythm and “good endings”: Effects of temporal structure on tonality judgments. *Perception & Psychophysics, 46*(1), 9-17. doi:https://www.doi.org/10.3758/BF03208069

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics, 53*(6), 585-600. doi:https://www.doi.org/10.3758/BF03211736

Bouwer, F. L., Werner, C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia, 85*, 80-90. doi:https://www.doi.org/10.1016/j.neuropsychologia.2016.02.018

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience, 20*(2), 226-239. doi:https://www.doi.org/10.1162/jocn.2008.20018

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 869-883. doi:https://www.doi.org/10.1037/0096-1523.7.4.869

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron, 88*(1), 2-19. doi:https://www.doi.org/10.1016/j.neuron.2015.09.019

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience, 7*(68), 1-16. doi:https://www.doi.org/10.3389/fnsys.2013.00068

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience, 35*(45), 15187-15198. doi:https://www.doi.org/10.1523/JNEUROSCI.2397-15.2015

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience, 32*, 1791-1802. doi:https://www.doi.org/10.1523/JNEUROSCI.4107-11.2012

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation.

Neuropsychologia, 38(11), 1531-1549. doi:https://www.doi.org/10.1016/S0028-3932(00)00053-1

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage, 47*(2), 735-744. doi:https://www.doi.org/10.1016/j.neuroimage.2009.04.090

Jiang, C., Liu, F., & Wong, P. C. (2017). Sensitivity to musical emotion is influenced by tonal structure in congenital amusia. *Scientific Reports, 7*(1), 7624. doi:https://www.doi.org/10.1038/s41598-017-08005-x

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323-355. doi:https://www.doi.org/10.1037/0033-295X.83.5.323

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*(6), 621-634. doi:https://www.doi.org/10.3758/BF03210494

Jones, M. R. (2009). Musical time. *The handbook of music psychology* (pp. 81-92).

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459-491. doi:https://www.doi.org/10.1037//0033-295X.96.3.459

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53*(1), 59-96. doi:https://www.doi.org/10.1016/j.cogpsych.2006.01.003

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313-319. doi:https://www.doi.org/10.1111/1467-9280.00458

Koelsch, S. (2013). *Brain and Music*. Oxford, UK: John Wiley & Sons.

Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience, 12*(3), 520-541. doi:https://www.doi.org/10.1162/089892900562183

Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain Research, 1212*, 55-62. doi:https://www.doi.org/10.1016/j.brainres.2007.10.078

Koelsch, S., & Jentschke, S. (2010). Differences in electric brain responses to melodies and chords. *Journal of Cognitive Neuroscience, 22*(10), 2251-2262. doi:https://www.doi.org/10.1162/jocn.2009.21338

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology, 44*(3), 476–490. doi:https://www.doi.org/10.1111/j.1469-8986.2007.00517.x

Koelsch, S., & Mulder, J. (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology, 113*(6), 862-869. doi:https://www.doi.org/10.1016/S1388-2457(02)00050-0

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America, 110*(38), 15443-15448. doi:https://www.doi.org/10.1073/pnas.1300272110

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology, 39*(5), 657–663. doi:https://www.doi.org/10.1017.S0048577202010508

Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General, 121*(4), 459-479. doi:https://www.doi.org/10.1037/0096-3445.121.4.459

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Kutas, M., & Kluender, R. (1994). What is who violating? A reconsideration of linguistic violations in light of event-related brain potentials. *Cognitive electrophysiology* (pp. 183-210). Birkhäuser, Boston: MA.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences, 1169*(1), 46-57. doi:https://www.doi.org/10.1111/j.1749-6632.2009.04550.x

Lebrun-Guillaud, G., & Tillmann, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception & Psychophysics, 69*(8), 1450-1459. doi:https://www.doi.org/10.3758/BF03192959

Loui, P., Grent, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research, 25*(3), 678-687. doi:https://www.doi.org/10.1016/j.cogbrainres.2005.08.019

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146-157. doi:https://www.doi.org/10.1111/psyp.12639

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience, 23*(9), 2252-2267. doi:https://www.doi.org/10.1162/jocn.2010.21542

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*, 177-190. doi:https://www.doi.org/10.1016/j.jneumeth.2007.03.024

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage, 38*(2), 331-345. doi:https://www.doi.org/10.1016/j.neuroimage.2007.07.034

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*, 156869. doi:https://www.doi.org/10.1155/2011/156869

Patel, A. D. (1998). Syntactic processing in language and music: Different cognitive operations, similar neural resources? *Music Perception, 16*(1), 27-42. doi:https://www.doi.org/10.2307/40285775

Patel, A. D. (2008). *Music, Language, and the Brain*. New York: Oxford University Press.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57. doi:https://www.doi.org/10.3389/fnsys.2014.00057

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience, 18*(9), 1545-1554. doi:https://www.doi.org/10.1162/jocn.2006.18.9.1545

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task and pitch structure on pitch-time interactions in music. *Memory & Cognition, 37*(3), 368-381. doi:https://www.doi.org/10.3758/MC.37.3.368

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception* *and Performance, 35*(5), 1598-1617. doi:https://www.doi.org/10.1037/a0016456

Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Perception, 26*(5), 401-413. doi:https://www.doi.org/10.1525/mp.2009.26.5.401

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992. doi:https://www.doi.org/10.3758/BF03206433

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review, 20*(3), 403-452. doi:https://www.doi.org/10.3758/s13423-012-0371-2

Russo, F. A., Thompson, W. F., & Cuddy, L. L. (2015). Effects of emergent-level structure on melodic processing difficulty. *Music Perception, 33*, 96-109. doi:https://www.doi.org/10.1525/mp.2015.33.1.96

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341-349. doi:https://www.doi.org/10.1111/1469-8986.00172.x

Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review, 79*, 369-382. doi:https://www.doi.org/10.1037/h0033118

Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex, 18*(5), 1169-1178. doi:https://www.doi.org/10.1093/cercor/bhm149

Sun, L., Liu, F., Zhou, L., & Jiang, C. (2018). Musical training modulates the early but not the late stage of rhythmic syntactic processing. *Psychophysiology, 55*(2), e12983. doi:https://www.doi.org/10.1111/psyp.12983

Tillman, C. M., & Wiens, S. Behavioral and ERP indices of response conflict in Stroop and flanker tasks. *Psychophysiology, 48*(10), 1405-1411. doi:https://www.doi.org/10.1111/j.1469-8986.2011.01203.x

Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research, 70*(5), 345-358. doi:https://www.doi.org/10.1007/s00426-005-0222-0

Trainor, L. J., & Hannon, E. E. (2013). Musical development. In D. Deutsch (Ed.), *The Psychology of Music* (3 ed., pp. 423-498). London: Elsevier.

Wang, L., Zhu, Z., Bastiaansen, M., Hagoort, P., & Yang, Y. (2013). Recognizing the emotional valence of names: An ERP study. *Brain and language, 125*(1), 118-127. doi:https://www.doi.org/10.1016/j.bandl.2013.01.006

Wong, P. C., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception, 27*(2), 81-88. doi:https://www.doi.org/10.1525/mp.2009.27.2.81

Zhang, Q., Guo, C. Y., Ding, J. H., & Wang, Z. Y. (2006). Concreteness effects in the processing of Chinese words. *Brain and language, 96*(1), 59-68. doi:https://www.doi.org/10.1016/j.bandl.2005.04.004

Zhou, L., Liu, F., Jiang, J., Jiang, H., & Jiang, C. (2019). Abnormal neural responses to harmonic syntactic structures in congenital amusia. *Psychophysiology, 56*(9), e13394. doi:https://www.doi.org/10.1111/psyp.13394

Acknowledgements

We thank Dr. Danchao Cai for her help with our data analysis. This work was supported by National Natural Science Foundation of China Grant No. 31470972 (CJ and FL), Grant No. 31500876 (LZ), the Australian Research Council Discovery Grant, DP130101084 (WT), and European Research Council (ERC) Starting Grant ERC-StG-2015, CAASD, 678733 (FL and CJ). None of the authors has a conflict of interest.

Figure Captions

Figure 1. Samples of the musical sequences used in the formal experiment. There were four conditions: (a) regular metric and harmonic structures, (b) regular metric and irregular harmonic structures, (c) irregular metric and regular harmonic structures, (d) irregular metric and harmonic structures

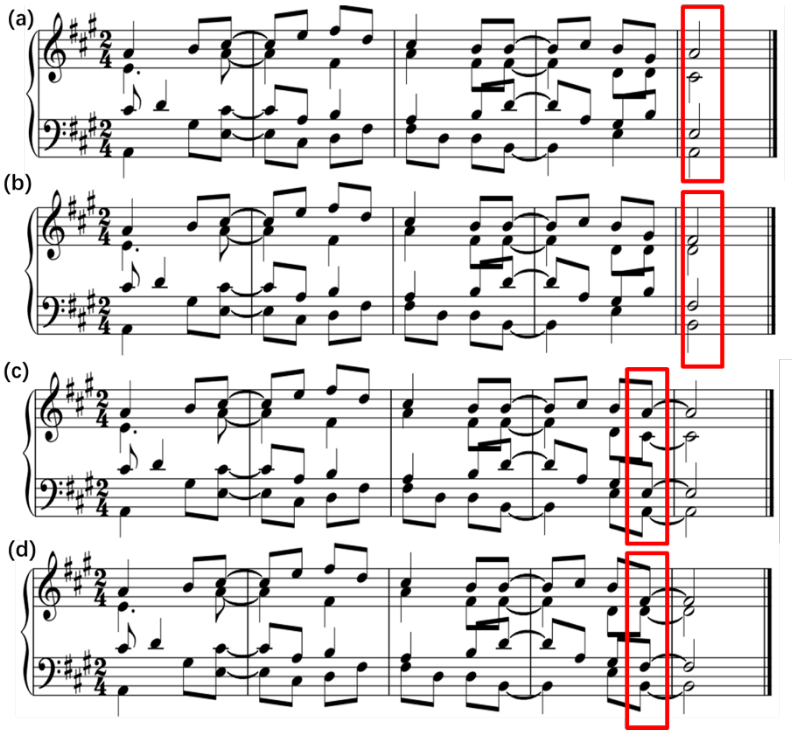
Figure 2. Mean ratings of the degree of completeness of the sequences with four types of endings.

Figure 3. Scalp distributions of the irregular-minus-regular metric difference waves in the 145–229 msec latency range and grand mean ERP waveforms elicited by regular and irregular metric endings collapsed over the harmonic conditions.

Figure 4. The difference waveforms of irregular-minus-regular metric endings under regular and irregular harmonic conditions, and the scalp distributions of the two difference waves in the 515–800 msec latency range.

Figure 5. Grand mean ERP waveforms elicited by regular and irregular harmonic endings at four electrode sites in the metric regular and irregular conditions. Gray-shaded areas indicate significant time windows. Scalp distributions of the irregular-minus-regular harmonic difference waves in the 300–647 msec latency range in the metric regular condition and in the 335–774 msec latency range in the metric irregular condition.

Figure 6. Grand mean ERP waveforms elicited by regular and irregular metric endings at four electrode sites in the harmonic regular and irregular conditions. Gray-shaded areas indicate significant time windows. Scalp distributions of the irregular-minus-regular metric difference waves in the 300–800 msec latency range in the harmonic regular condition and in the 335–774 msec latency range in the harmonic irregular condition.

Figure 1

Regular metric and harmonic structure (regular)

Regular metric and irregular harmonic structure (single violation)

Irregular metric and regular harmonic structure (single violation)

Irregular metric and harmonic structure (double violation)

Figure 2

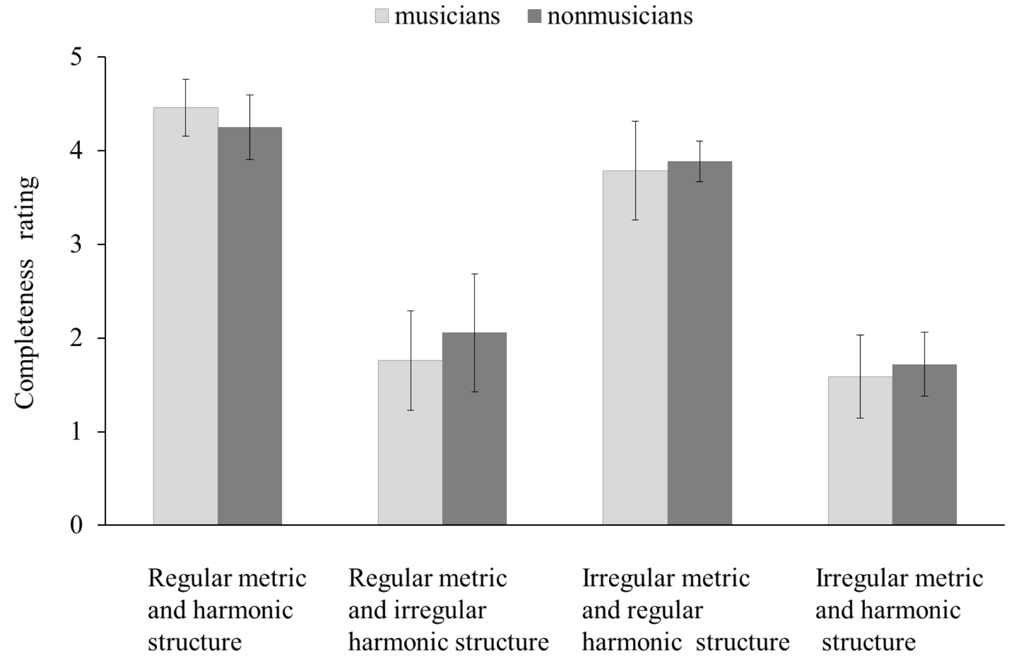

musicians
nonmusicians
Completeness rating
5
4
3
2
1
0
Regular metric and harmonic structure
Regular metric and irregular harmonic structure
Irregular metric and regular harmonic structure
Irregular metric and harmonic structure

Figure 3

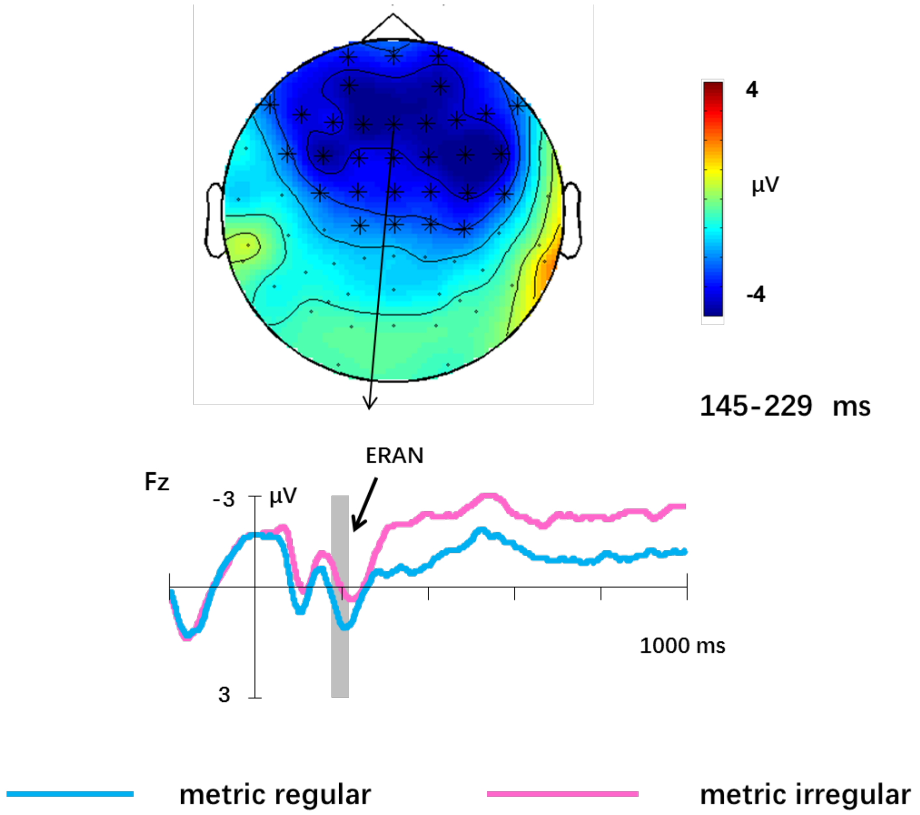

4
μV
-4
145-229 ms
ERAN
Fz
-3
μV
3
1000 ms
metric regular
metric irregular

Figure 4

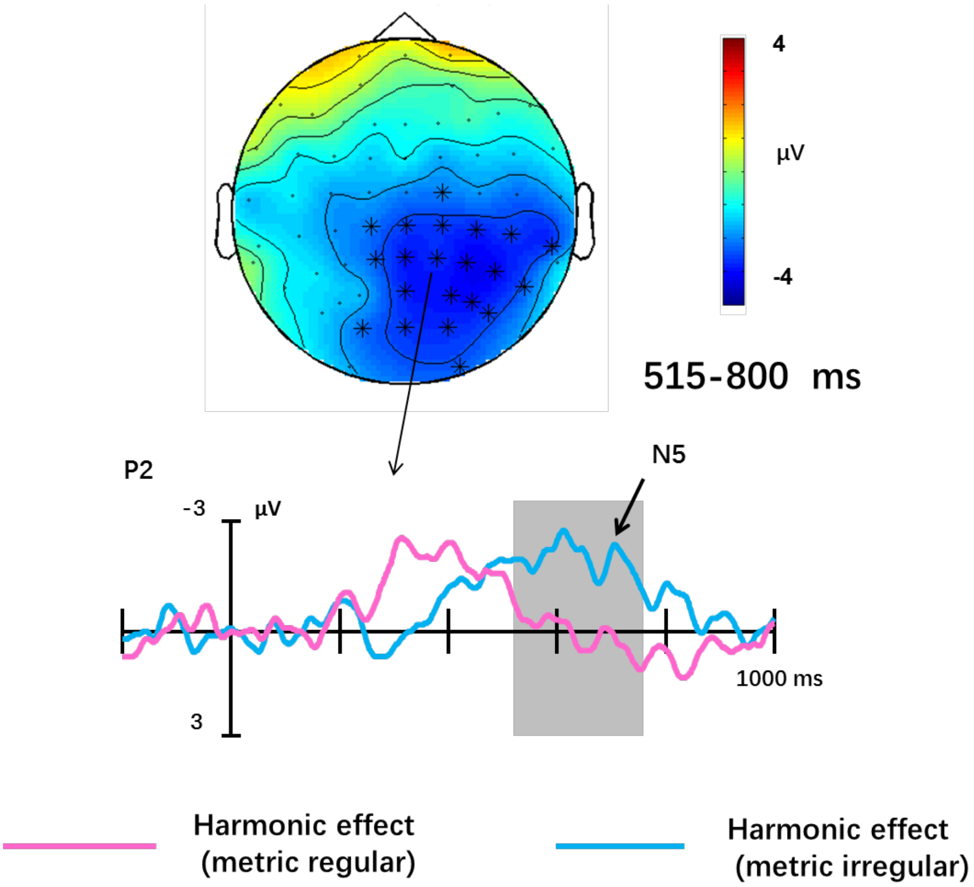

4
μV
-4
515-800 ms
P2
N5
-3
μV
3
1000 ms
Harmonic effect (metric regular)
Harmonic effect (metric irregular)

Figure 5

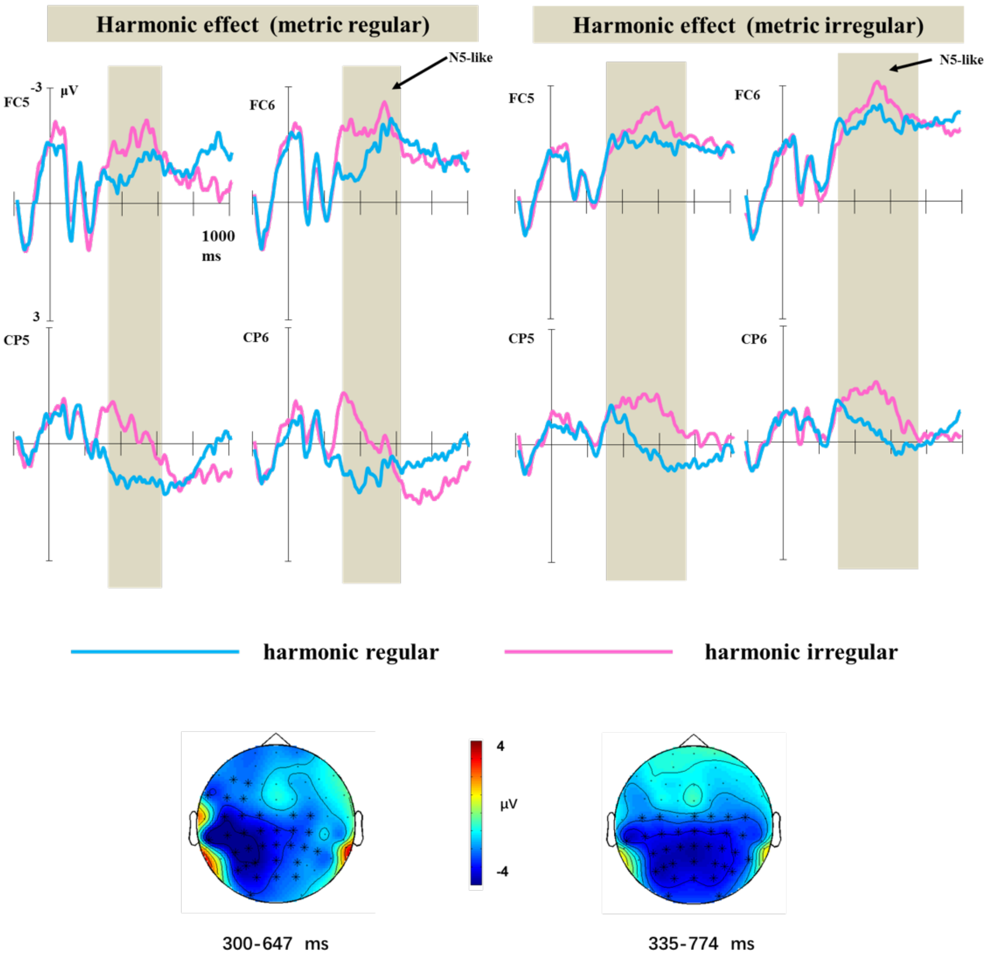
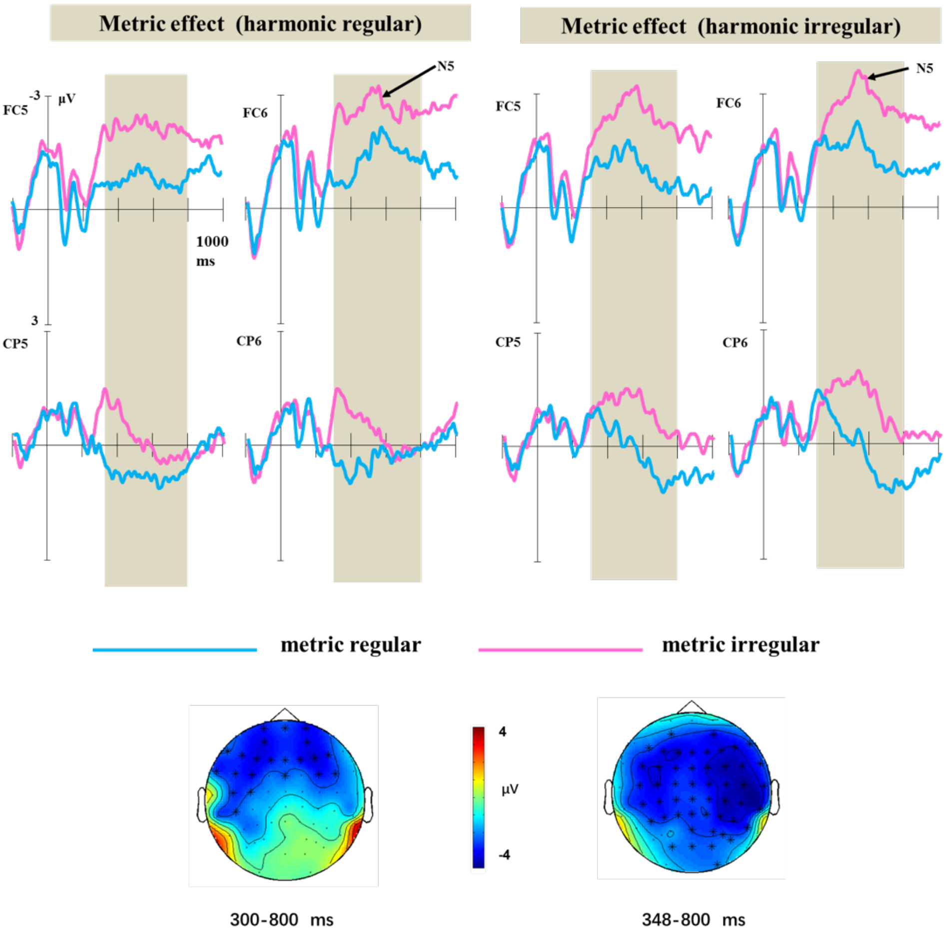

Harmonic effect (metric regular)
Harmonic effect (metric irregular)
N5-like
N5-like
FC5
-3
µV
FC6
1000
ms
3
FC5
FC6
CP5
CP6
CP5
CP6
harmonic regular
harmonic irregular
4
µV
-4
300-647 ms
335-774 ms

Figure 6


Metric effect (harmonic regular)
Metric effect (harmonic irregular)
-3
μV
3
FC5
FC6
CP5
CP6
N5
1000 ms
metric regular
metric irregular
4
μV
-4
300-800 ms
348-800 ms