Neuropsychologia 121 (2018) 164–174

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Syntactic complexity and musical proficiency modulate neural processing of non-native music

Xie Ma[a,b,c], Nai Ding[d,e,f], Yun Tao[b,c], Yu Fang Yang[a,*]

[a] Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[b] College of Educational Science and Management, Yunnan Normal University, Kunming, China
[c] Key Laboratory of Educational Informatization for Nationalities, Yunnan Normal University, Kunming, China
[d] College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou, China
[e] Key Laboratory for Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou, China
[f] State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China

ARTICLE INFO

Keywords:
Syntax
Non-native music
Finite state grammar
Phrase structure grammar
EEG

ABSTRACT

In music, chords are organized into hierarchical structures on the basis of musical syntax and the syntax of Western music can be implicitly acquired by listeners growing up in a Western musical culture. Here, we investigated whether Western musical syntax of different complexities can be implicitly acquired by non-native listeners growing up in China. This study used electroencephalography (EEG) to measure how the neural responses to musical sequences that either follow a simple rule, i.e., finite state grammar (FSG), or a complex rule, i.e., phrase structure grammar (PSG), are affected. We tested three groups of Chinese listeners who varied in their proficiency and experience in Western music. Only the high-proficiency group had received formal Western musical training, whereas the low- and moderate-proficiency groups varied in their degree of exposure to Western music. The results showed that in the FSG condition, the event-related potentials (ERPs) evoked by regular and irregular final chords were not significantly different in the low-proficiency group. In contrast, in the moderate- and high-proficiency groups, the irregular final chords evoked an ERAN-N5 biphasic response. In the PSG condition, however, only the high-proficiency group showed an ERAN-N5 biphasic response evoked by irregular final chords. This study provides evidence that although simple structures of Western music, such as FSG, can be acquired by long-term implicit learning, the acquisition of more complex structures, such as PSG, merely from exposure to western music may not be as easy.

1. Introduction

The Western music system features syntactic structures based on tonal harmony relationships. According to the generative theory of tonal music (GTTM) (Lerdahl and Jackendoff, 1987) and the generative syntax model (GSM) (Rohrmeier, 2011); the hierarchical organizational structures in music may be similar to those found in language. A simple form of syntax is finite state grammar (FSG), which assumes a linear structure that can be fully specified by the transition probability between a finite numbers of states (e.g., chords). Another type of grammar is referred to as phrase structure grammar (PSG), hierarchical grammar, or context-free grammar, which permits hierarchically organized structures formed by center-embedding (Rohrmeier, 2011; Rohrmeier et al., 2014; Rohrmeier and Rebuschat, 2012). For instance, in the language sequence “The teacher [that the boy saw] was tall”, the clause “the boy saw” is center-embedded into the main sentence “The teacher was tall”. Similarly, in the musical sentence “key 1–key 2–key 1”, “key 2” can be regarded as being center-embedded into a key 1 structure (Rohrmeier and Cross, 2009; Woolhouse et al., 2016). To process such hierarchical structures, learners need to integrate the separated (or non-adjacent) “key 1” structures’ harmonic relationship (Koelsch et al., 2013; Rohrmeier and Cross, 2009; Woolhouse et al., 2016). FSG and PSG structures differ significantly in their abstraction and complexity (Rohrmeier, 2011; Rohrmeier and Cross, 2009). Studies in the field of language and other disciplines have shown that compared to FSG, the processing of PSG usually demands more cognitive control resources, such as working memory, attention, and inhibition (Badre, 2008; Jeon and Friederici, 2015; Koechlin and Summerfield, 2007; Makuuchi et al., 2013).

Within the Western musical culture, a number of studies have reported that the perception of tonal syntax is not significantly related to their listeners’ experience, suggesting that harmonic structure can be

* Corresponding author.
E-mail address: yangyf@psych.ac.cn (Y.F. Yang).

https://doi.org/10.1016/j.neuropsychologia.2018.10.005
Received 24 March 2018; Received in revised form 21 September 2018; Accepted 8 October 2018
Available online 22 October 2018
0028-3932/

acquired without formal musical training for both FSG (Bigand et al., 1999, 2014; Bigand and Pineau, 1997; Koelsch et al., 2000, 2007) and, the more complex, PSG (Koelsch et al., 2013; Rohrmeier and Cross, 2009; Woolhouse et al., 2016). EEG studies of the processing of tonal syntax have focused on two ERP components elicited by musical syntactical irregularities. The early anterior negativity (ERAN) reflects musical syntactic processing and occurs roughly within the 150–300 ms time window (Koelsch and Jentschke, 2009; Koelsch et al., 2013; Przysinda et al., 2017). N5 is a late negative wave reflecting harmonic integration, and it roughly occurs within the time window of 400–850 ms (Koelsch et al., 2000, 2007, 2013; Koelsch and Jentschke, 2009; Zhang et al., 2017). For two-year-old children, the ERAN appears in FSG on-line processing, indicating that the FSG in music can be acquired through implicit learning (such as statistical learning), in addition, it does not need a long time for the implicit experiences to accumulate (Jentschke et al., 2014). For PSG processing, the ERAN and N5 could be evoked by a mismatch in tonality for both musicians and non-musicians (Koelsch et al., 2013), suggesting that native learners can also acquire PSG through implicit learning.

Research of musical syntax cognition has mostly focused on Western listeners. It remains controversial, however, whether the syntactic cognition in music is across cultures and whether listeners are sensitive to the syntactic structures in their non-native musical cultures. Growing evidence from studies on non-native listeners has revealed that cultural bias influences music cognition (Mcdermott et al., 2016; Patel and Demorest, 2013) and listeners can process music from their culture (i.e., native music) better than music from other cultures (i.e., non-native music) (Demorest et al., 2016; Patel and Demorest, 2013). Even when non-native listeners were fully exposed to a non-native music system, they still showed lower musical competence relative to native listeners (Demorest et al., 2016; Demorest and Osterhout, 2012). Furthermore, it was found that the acquisition of non-native music involves more complicated cognitive and neural mechanisms (Demorest et al., 2010; Nan et al., 2008).

A few researchers have investigated the processing of non-native music syntax (Akrami and Moghimi, 2017; Brattico et al., 2013; Demorest and Osterhout, 2012; Eerola et al., 2009; Krumhansl et al., 2000; Wong et al., 2009). Some found that the behavioral performance on musical syntactic tasks are similar for native and non-native listeners (Eerola et al., 2009; Krumhansl et al., 2000), whereas others found poorer performance or distinct brain responses for non-native listeners relative to native listeners (Brattico et al., 2013; Demorest and Osterhout, 2012; Wong et al., 2009). These studies concentrated on local or simple musical structure for which music expectancy may be more likely to be generated on the basis of sensory priming (Bigand et al., 1999, 2014; Collins et al., 2014). Thus far, however, little is known about whether people can perceive more abstract and complex structures (Akrami and Moghimi, 2017; Rohrmeier and Rebuschat, 2012), i.e., non-adjacent and nested structures, in non-native music, and especially whether the structure complexity will affect the perception and acquisition of non-native music.

Research on non-native music processing could be compared with research on non-native language processing. In bilingual studies, it has been shown that non-native learners usually have difficulty grasping syntax (Clahsen and Felser, 2006; Foucart and Frenck-Mestre, 2012; Hahne and Friederici, 2001). More importantly, the deficits in syntactic processing are more obvious for complicated syntactic structures (e.g., passive sentences, fill-gap sentences) (Clahsen and Felser, 2006; Felser and Roberts, 2007; Kim et al., 2015). Even some non-native learners who can effectively process simple structures (e.g., adjacent structure), show difficulty integrating complex syntactic structures (e.g., non-adjacent structure) (Bañón et al., 2012; Dowens et al., 2010; Rossi et al., 2006). Furthermore, neuroimaging studies have shown that native and non-native learners show greater differences in cortical activation when processing more complex syntactic structures (Hasegawa et al., 2002; Yokoyama et al., 2006). Non-native learners' deficit in syntactic processing has been attributed to missing the critical period of learning (DeKeyser and Larson-Hall, 2005) and inhibiting the inference of native language (Foucart and Frenck-Mestre, 2011; Kim et al., 2015). In general, non-native learners can hardly process syntactic structure as automatically and as effectively as native speakers and have to rely more on general cognitive resources (Clahsen and Felser, 2006; Dowens et al., 2010; Mcdonald, 2006; Suh et al., 2007; Zhou et al., 2017). If the poorer syntactic processing performance in non-native learners is related to limitations in general cognitive resources, it naturally deteriorates for more complex syntactic structures that have a stronger demand for cognitive control (Bañón et al., 2012; Clahsen and Felser, 2006; Hasegawa et al., 2002; Suh et al., 2007; Yokoyama et al., 2006).

Several similarities between the hierarchical organization of musical and linguistic syntax have been proposed (Lerdahl and Jackendoff, 1987; Rohrmeier, 2011). The neural mechanisms underlying musical syntax parsing may parallel those underlying linguistic syntax processing (Patel, 2003). Furthermore, cognitive control may play an especially pivotal role in constructing deep and abstract syntactic representations for both music and language (Farbood et al., 2015; Fedorenko et al., 2009; Slevc and Okada, 2015). More importantly, it has been shown that some adverse factors impeding non-native language processing also exist in non-native music processing, for example, *the non-native musical syntax processing might be interfered with by the native musical syntactic schemas, and the more experience of native-music, more possibly hampered the perception of non-native music, especially when the syntax of two musical syntax are very different from each other (*Brattico et al., 2013; Matsunaga et al., 2012; Wong et al., 2011*)*. Thus the processing of non-native musical syntax may need more cognitive resource to inhibit the activation of native musical syntactic representation (Clahsen and Felser, 2006).

An individual's proficiency and experience in non-native music also has a considerable effect on non-native music processing. Some studies have found that listeners with more training did not perform differently in cross-cultural music perception tasks compared with those who had no training (Demorest et al., 2008). Other studies, however, report that increasing exposure to and training in non-native music can significantly promote non-native music perception (Wong et al., 2009; Zhang et al., 2016). In the language domain, several studies address how language proficiency affects syntactic processing in non-native learners (Dowens et al., 2010; Morganshort et al., 2012; Rossi et al., 2006). Furthermore, it was revealed that proficiency, syntactic complexity, and the difference between native and non-native syntax all affect syntactic processing in non-native listeners (Dowens et al., 2010; Jeon and Friederici, 2013, 2015; Rossi et al., 2006). How the multiple factors interact, however, remains uncertain.

In this study, we investigate whether the structural complexity and musical proficiency modulate neural processing of non-native music. In particular, we focus on how Chinese listeners process Western tonal music, as the Chinese and Western musical cultures have clear differences. A key feature of tonal syntax distinguishing Chinese and Western musical cultures is the "musical diatonic scale." Traditional Chinese scales are mainly composed of five-tone scales (i.e., pentatonic scales); whereas Western scales primarily comprise seven-tone scales. A number of Chinese composers in modern times have also heavily relied on pentatonic scales in music creation (Fan, 2003; Li, 2001; Xu, 2014). Furthermore, compared with Western music, the expression of Chinese music (even in popular music) is conveyed mainly by melodic rather than harmonic structure (Xu, 2014; Rohrmeier et al., 2014). Chinese listeners are accustomed to the dominant Chinese tonal music culture but vary significantly in their proficiency and experience in Western tonal music. Some receive formal Western musical training; some listen to Western music frequently; whereas others have little exposure to Western music.

Two musical sequences differing in structure complexity are used in this study: a simple FSG sequence and a complex PSG sequence in Western tonal music. Additionally, three groups of subjects varying in their proficiency with Western music were recruited. In sum, the study will examine how listeners with different levels of musical proficiency process simple and complex musical syntax from a non-native musical culture, and whether musical proficiency interacts with the complexity of musical syntax.

2. Methods

2.1. Participants

All participants were born and grew up in China (age-range, 20–24 years; M = 22.57; SD = 1.22). Among them, 17 had high proficiency in Western tonal music (nine female, eight male), 16 had moderate proficiency (eight female, eight male), and 16 had low proficiency (seven female, nine male). All participants were students from Yunnan Normal University and Yunnan Arts University. The low-proficiency level group received no professional music training. Furthermore, they had no habit or preference of listening to Western music, but were occasionally passively exposed to Western music through TV shows, internet media, and other media. The moderate-proficiency level group received no formal music training; however, they preferred to listen to Western music for at least 2 h each week, including Western pop and classical music. The high-proficiency group received more than 7 years of professional instruction and training, playing Western instruments such as the piano and violin. In addition, we controlled for experience with Chinese folk music because a listener's formal training in native music may affect the perception of non-native music (Bañón et al., 2012; Matsunaga et al., 2012). In all groups, no one reported any history of being full-time students studying traditional Chinese music in musical academies or any experience performing traditional Chinese music as a main daily activity. All participants had no abnormal hearing or absolute pitch and provided signed informed consent for study participation. This study conforms to the ethical principles of the Declaration of Helsinki (World Medical Organization, 1996).

2.2. Stimuli

The PSG sequences were based on those used in the study by Koelsch et al. (2013). The first chorale was selected from BWV 302 by J. S. Bach, while the second chorale was from BWV 373 by J. S. Bach. The harmonic structure of these two chorales was complex; they featured a segment of lower level in a local key, embedded in a segment of higher level in the overarching key. As illustrated in Figs. 1C and S1C. The first phrase ended with the dominant; the second phrase then progressed to another key (referred to as a transition or modulation) and concluded with the initial tonic through an authentic cadence. From the GTTM and GSM perspective (Lerdahl and Jackendoff, 1987; Rohrmeier, 2011), the final chord of the two chorales hierarchically prolonged the first chord of the first phase and closed the dominant of the first phrase ending. These hierarchical relationships were nested and center-embedded (Rohrmeier, 2011; Rohrmeier et al., 2014; Rohrmeier and Rebuschat, 2012). Therefore, the harmonic sequence with a PSG structure is similar to the English sentence "The teacher [that the boy saw] persistently tries". Because this study is mainly concerned with tonal structure processing, to remove the interference from rhythm, the main harmonic structure (function) was extracted (reduction) with a uniform rhythm using quarter notes based on former similar studies (Rohrmeier and Cross, 2009; Woolhouse et al., 2016). The irregular PSG sequences were created on the basis of the regular sequences (see Fig. 1D), with the first phrase being transposed up a major second in the first chorale and transposed down a fourth in the second chorale. In this case, the final chord of the second phrase did not prolong the first chord nor close the dominant of the first phrase. Importantly, the second phrase of the regular and the irregular sequences were identical. Therefore, the manipulation only led to an irregular hierarchical dependency.

The FSG sequence was obtained by modifying the PSG sequences using a method adopted by previous studies (Bigand et al., 1999; Rohrmeier and Cross, 2009; Woolhouse et al., 2016) (see Fig. 1A). Long-distance dependency in the PSG sequence imposes additional memory load for sequence processing (Makuuchi et al., 2009, 2013). To factor out this effect while reducing the effect of sensory priming (Bigand et al., 2014; Collins et al., 2014), we also created non-adjacent dependency in the FSG sequence (Bigand et al., 1999), for example by detaching the dominant and tonic chords. The first phrase remained unchanged and ended with a half cadence on the dominant. The second phrase continued in the same key but did not begin with the tonic. The ending of the second phrase was still at the initial tonic, which also remained unchanged (cf. the PSG sequence; see Fig. 1). In this sequence, each segment is in a single key. The harmonic structure with FSG is similar to the FSG structure in the English sentence "The teacher [in a blue coat] persistently tries." An irregular FSG sequence was created on the basis of each regular FSG sequence (see Fig. 1B). Similar to the procedure used for the irregular PSG, the first phrase was transposed up a major second and transposed down a fourth in the first and second chorale respectively. In this case, the final chord at the end of the second phrase did not close the dominant established by the first phrase.

Although the transition between the last chord of the first phrase and the first chord of the second phrase (see the dominant with the fermata in Figs. 1 and S1) was plausible in both regular and irregular versions; the local transition probabilities were lower in the irregular versions than in the regular versions (Koelsch et al., 2013).

Using Sibelius 6.2 software, each stimulus at each condition was transposed into six different major keys, after which the files were rendered using a piano sound at a tempo of 100 beats per minute and archived in wav format. The rendering processes yielded the following stimuli: 6 keys × 2 chorales × 2 sequence type (regular, irregular) × 2 syntax types (FSG, PSG), yielding 48 stimuli in total. Additionally, to assess a focus of attention for participants' listening, 32 stimuli were created for the timbre detection task, which required the subjects to monitor the timbre of the musical stimuli. These stimuli modified, wherein one chord of the chorale was played using a bassoon instead of a piano. These timbre deviants were spread equally at four different positions and for eight different conditions (2 chorales × 2 sequence types × 2 syntax types).

2.3. Procedure

The EEG (electroencephalographic) recording session involved the participants watching a muted video without captions while playing the stimuli at 60 dB SPL (sound pressure level). All 48 stimuli were presented five times in a random order, mixed with the 32 timbre variants. The task for the subjects was to monitor the timbre of the musical stimuli, keep track of the stimuli, and identify the timbre deviants by clicking a response button. The 32 stimuli with timbre deviants were not used in the analysis of ERPs. A total of 272 (48 × 5 + 32) stimuli were played for an experimental duration of approximately 70 min. After the recording, the participants were tasked to assess (conclusiveness rating) 16 stimuli (2 keys × 2 chorales × 2 syntax types × 2 sequence types), answering the question "How well did the last chord end the whole sequence?" on a scale of one (not at all) to nine (perfectly ended), followed by choosing their ground of their conclusiveness from four options: (i) guessing, (ii) their intuition, (iii) knowing the rule, or (iv) knowing the piece.

2.4. ERP recordings and data analysis

Using Brain Products (Munich, Germany), brain electrical activity was recorded from 64 Ag/AgCl electrodes mounted on an elastic cap. The data were referenced online to FCz and offline re-referenced to the algebraic average of the left and right mastoids (Luck, 2014). An

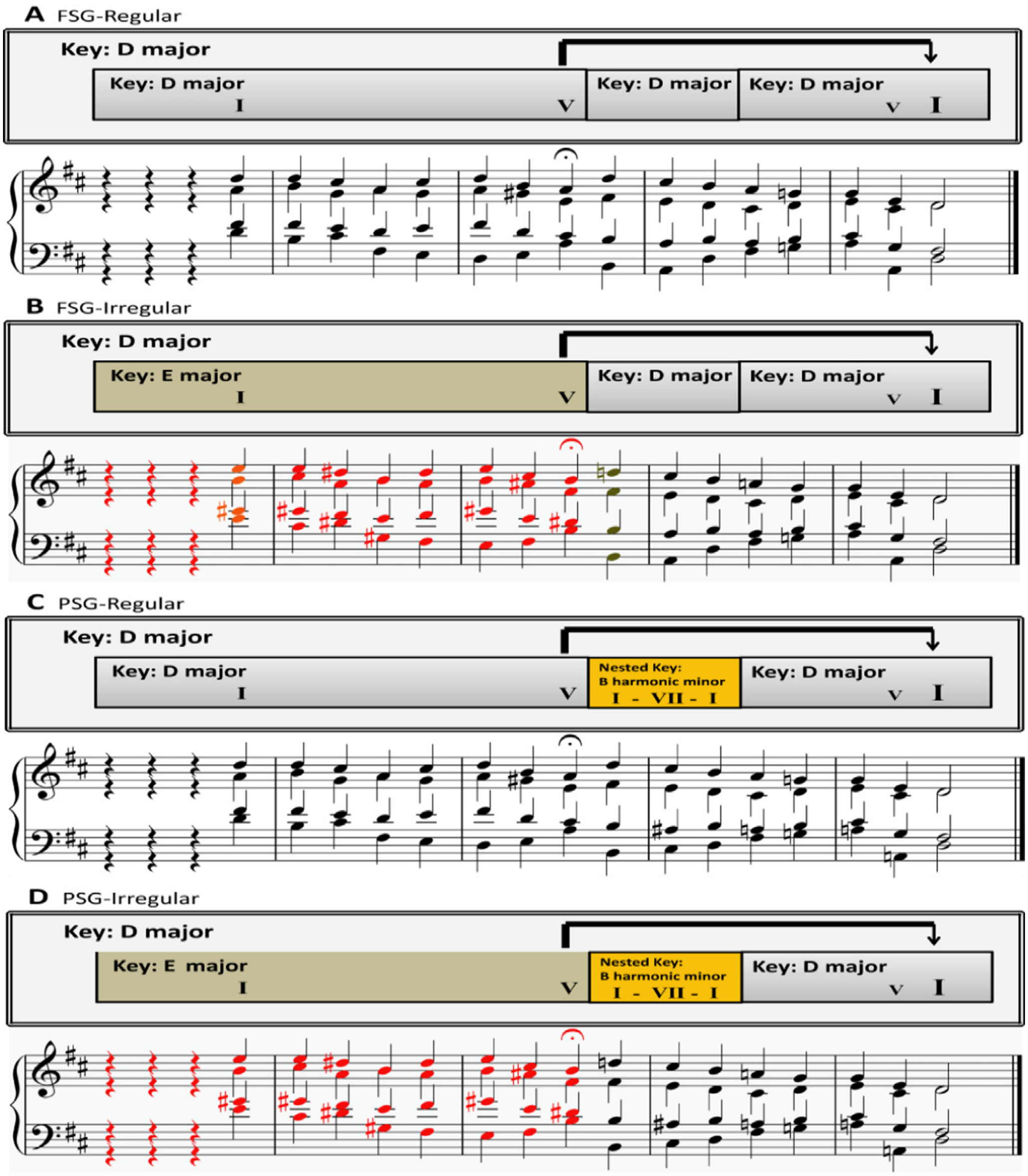

Fig. 1. Illustration of four experimental conditions (regular and irregular versions of the FSG and PSG sequences). (A, B) The FSG sequences do not contain any center-embedded structure. The tonal relationship between the initial segment and final chord is closed for the regular FSG sequence (both are in D major) but less closed for the irregular FSG sequences (the initial key is E major; the final chord as the authentic cadence is closed is in D major). (C, D) The PSG sequences have a center-embedded structure; a B harmonic minor segment is inserted in an overarching key of D major. The tonal relationship between the initial segment and the final chord is closed (both are in D major) for the regular PSG sequence but less closed for the irregular PSG sequence (the initial key is E major; the final chord as the authentic cadence is closed is in D major).

impedance of less than 5 kΩ was maintained in all electrodes. The horizontal electrooculograms and vertical electrooculograms were collected from the left against the right orbital rim and infra and supra arbitrarily at the left eye, respectively. Then continuous sampling at 1000 Hz and an FIR filter (0.05–100 Hz band filter) to amplify the signals for offline analysis were performed.

The pre-processing of raw EEG data was performed using Brain Vision Analyzer 1.05, part of Brain Products. At first, ocular artifacts were corrected and data were filtered offline with a band-pass filter of 0.1–25 Hz (24 dB/octave). The data were then segmented from 200 ms before to 1200 ms after the onset of each of the two chords of interest—the final chord and the first chord of the second phrase, with baseline correction from 200 ms to 0 ms preceding chord onset. Next, based on the Eye Movement Correction Algorithm (Gratton et al., 1983), eye movements and blinks (EOG artifacts) were corrected. To eliminate trials of poor quality, offline computerized artifact rejection (mean EOG voltage exceeding ± 80 μV) was used. In addition, trials containing artifacts caused by muscle movements or amplifier block were removed the same as trials containing peak-peak deflection more than ± 80 μV. Referring to that in the study by Koelsch (2013), the ERPs elicited by the two chords were marked and analyzed as (i) onset of the final chord (ii) onset of the first chord of the second phrase. For statistical analysis of ERPs, four regions of interest (ROIs) were computed: left anterior (F1, F3, F5, FC1, FC3, FC5), right anterior (F2, F4, F6, FC2, FC4, FC6), left posterior (CP1, CP3, CP5, P1, P3, P5), and right posterior (CP2, CP4, CP6, P2, P4, P6).

Global analyses of variance (ANOVAs) were conducted with the between-subjects factor of Western tonal musical proficiency group (low, moderate, high), and the within-subject factors of syntax type (FSG, PSG), sequence type (regular, irregular), hemisphere (left, right ROIs), and anterior–posterior distribution (anterior, posterior ROIs). The specific time windows for statistics based on the existing literature (Koelsch et al., 2000, 2007; Koelsch and Jentschke, 2009; Zhang et al., 2017), and preliminary analysis (we conducted a series of analyses in consecutive mean amplitude latency bins of 50 ms and found the exact initial and the closed time of the effects of sequence type). For the final chord, the preliminary analysis revealed that the early effects of sequence type begin with 150–200 ms window and end with 250–300 ms window in most conditions (at the different conditions of groups and syntax types); the later effects of sequence type begin with 400–450 ms window and end with 650–700 ms window in most conditions. So, the 150–250 ms and the 400–650 ms time windows were chosen for the statistical analysis of the final chord. For the first chord of the second phrase, the preliminary analysis revealed that the early effects of sequence type begin with 150–200 ms window and end with 250–300 ms window in most conditions; the N5 effects of sequence type begin with 500–550 ms window and end with 600–650 ms window in most conditions. So, the 150–250 ms and the 500–600 ms time windows were chosen for the statistical analysis of the first chord of the second phrase. Only the significant (marginally significant) effects of sequence type and significant interactions involving sequence type are reported. For all analyses, the degrees of freedom of the *F* ratio were corrected for violations of the sphericity assumption based on the Greenhouse-Geisser correction (Greenhouse and Geisser, 1959), and Bonferroni corrections were used for each comparison (Keppel, 1991).

3. Results

3.1. Behavioral data

During the EEG session, participants detected, on average, 97.94% (1.8) of the timbre-deviant chords [low: 98.20% (1.8%), moderate: 97.65% (2.1%), high: 97.94% (1.6%)]. Hit rates did not differ between groups, sequence type, or syntax type, indicating that the attentiveness of the participants was similar across all conditions.

For the conclusiveness ratings (see Fig. 2), a repeated-measure ANOVA with the between-subjects factor of group (low-, moderate-, and high-proficiency), and the within-subject factors of sequence type (regular, irregular) and syntax type (FSG, PSG) was carried out. The results indicated a main effect of sequence type [F (1, 46) = 19.52, $p < .001$, $\eta_p^2 = 0.30$] and an interaction effect among sequence type, syntax type, and group [F (1, 46) = 5.044, $p = .010$, $\eta_p^2 = 0.18$]. Simple effect analyses showed that for the low-proficiency group, no sequence type effect was found; for the moderate-proficiency group, the conclusiveness ratings obtained were higher for regular than irregular sequences only in the FSG condition ($p = .002$); however, for the high-proficiency group, the conclusiveness ratings were higher for regular than irregular sequences in both the FSG ($p < .001$) and PSG conditions ($p < .001$).

To further confirm the participants' Western music experience and proficiency, *we also assessed participants' awareness of their knowledge guiding conclusiveness ratings (see* Table 1*)*. A χ^2 test showed that individuals in the low-proficiency group chose “guessing” more frequently than those in the moderate- ($p = .007$) and high-proficiency ($p < .001$) groups; individuals in the high-proficiency group were more likely to choose “knowing the rule” than those in the moderate- ($p < .001$) and low-proficiency groups ($p < .001$); individuals in the moderate-proficiency group chose “intuition” more often than those in the high-proficiency group ($p = .002$). This result supports the fact that there were significantly different Western musical experiences proficiency among the three groups.

Table 1
The source attribution of conclusiveness ratings (%).

	Guessing	Intuition	Knowing the rule	Knowing the piece
LP Group	31.6	68.4	0	0
MP Group	19.5	79.3	0.4	0.8
HP Group	7	52.9	38.6	1.5

3.2. ERP data

3.2.1. Final chords

Figs. 3 and 4 illustrate the neuro-electrical responses to the final chords in different syntax conditions. In the FSG condition, the final chords of the irregular sequence induced an early frontal-central negativity that appeared at a standard ERAN latency range (around 150–250 ms after final chord onset) in the moderate- and high-proficiency groups, and a subsequent later frontal negativity that emerged at a standard N5 latency range (around 400–650 ms after final chord onset). However, neither of these effects was observed in the low-proficiency group. In the PSG condition, only for the high-proficiency group, the final chords of the irregular sequence induced an early frontal-central negativity that appeared in the 150–250 ms time window and a subsequent later frontal negativity that appeared in the 400–650 ms time window. However, neither of these effects not observed in the low- or moderate-proficiency groups.

3.2.1.1. Results in the 150–250 ms latency range. A repeated-measures ANOVA with the between-subjects factor of group (low-, moderate-, and high-proficiency), and within-subject factors of sequence type (regular, irregular), syntax type (FSG, PSG), hemisphere (left, right ROIs), and anterior–posterior distribution (anterior, posterior ROIs) was performed for a time window from 150 to 250 ms (the waves in the regular and irregular sequence are shown in Fig. 3 and 4; the amplitude values of the effects within these time windows are shown in Table 2). The results showed a main effect of sequence type [F (1, 46) = 5.13, $p = .028$, $\eta_p^2 = 0.10$] and a marginal interaction between sequence type, syntax type, anterior–posterior ROIs, and group [F (2, 46) = 3.13, $p = .053$, $\eta_p^2 = 0.12$]. The simple effect analyses showed that for the low-proficiency group, no sequence type effect was found; for the moderate-proficiency group, the negativity was larger for the irregular sequence than for the regular sequence in the FSG condition, and the effect was significant only in anterior regions ($p = .001$); for the high-

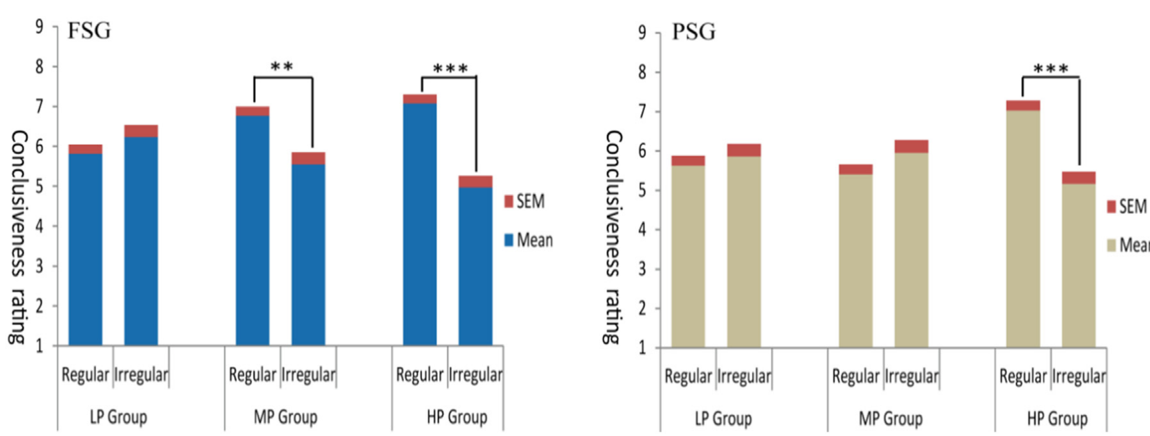

Fig. 2. Conclusiveness ratings on final chords. In the FSG condition (left), the mean conclusiveness ratings were higher for regular than for irregular sequences for both the moderate-proficiency ($p < .01$) and high-proficiency groups ($p < .001$). In the PSG condition (right), the mean conclusiveness ratings were higher for the regular than for irregular sequences only for the high-proficiency group ($p < .001$).

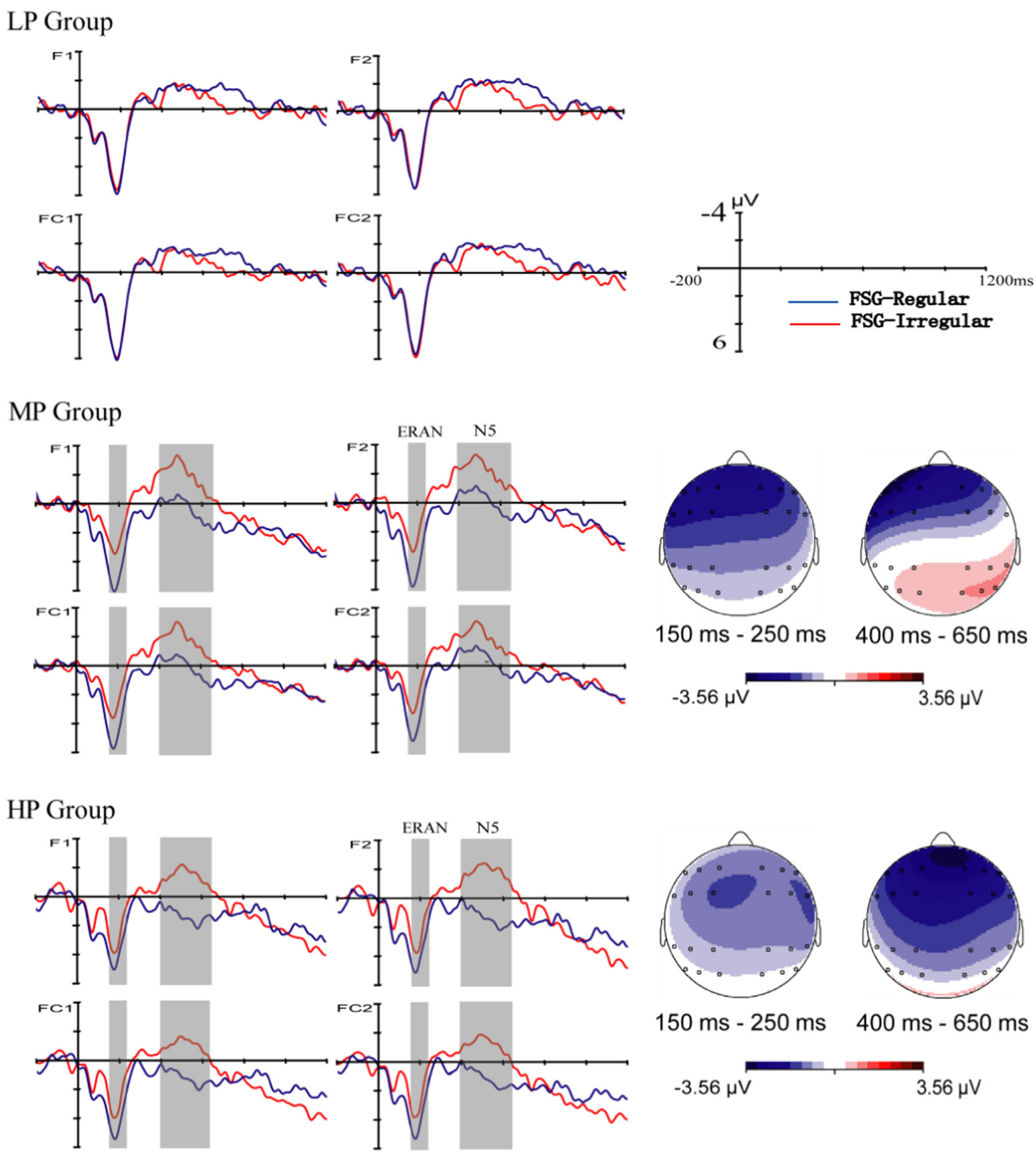

Fig. 3. ERP responses to the final chord in FSG sequences Event-related potentials (ERPs) evoked by the final chords in FSG sequences are shown separately for low-, moderate-, and high- proficiency groups. For the moderate- and high-proficiency groups, compared to the responses to regular sequences (blue waveforms), the responses to irregular sequences (red waveforms) evoked an early negativity (maximal around 150–250 ms) and a later negativity (maximal around 400–650 ms). Shaded areas indicate time windows used for the statistical analysis reported in the text. For the low-proficiency group, none of the ERP effects differed significantly between regular and irregular sequences. The right panel shows the scalp distribution of the early and late ERP effects, i.e., the difference potentials between regular and irregular sequences).

proficiency group, the negativity was larger for the irregular sequence than for the regular sequence both in the FSG and PSG condition, an effect that was significant only in the anterior regions for both the FSG (p = .032) and PSG conditions (p = .001).

3.2.1.2. Results in the 400–650 ms latency range. In the 400–650 ms time window, an analogy global ANOVA was carried out (the effects can also be seen in Figs. 3 and 4, and Table 2). The results revealed a main effect of sequence type [F (1, 46) = 8.66, p = .005, η_p^2 = 0.16] and an interaction among sequence type, syntax type, anterior–posterior ROIs, and group [F (2, 46) = 6.39, p = .004, η_p^2 = 0.22]. Simple effect analyses showed that for the low-proficiency group, no significant effect of sequence type was found; for the moderate-proficiency group, the negativity was larger for the irregular than regular sequence in the FSG condition, and the effect was mainly distributed in the anterior region (p = .001); for the high-proficiency group, the negativity was larger for the irregular than regular sequence in both the FSG and PSG conditions, and the effect was significant only in anterior regions in the FSG (p < .001) and PSG conditions (p < .001).

3.2.2. The first chords of the second phrase

The ERPs of the first chord of the second phrase show that the irregular sequences, compared to regular sequences, evoked an early

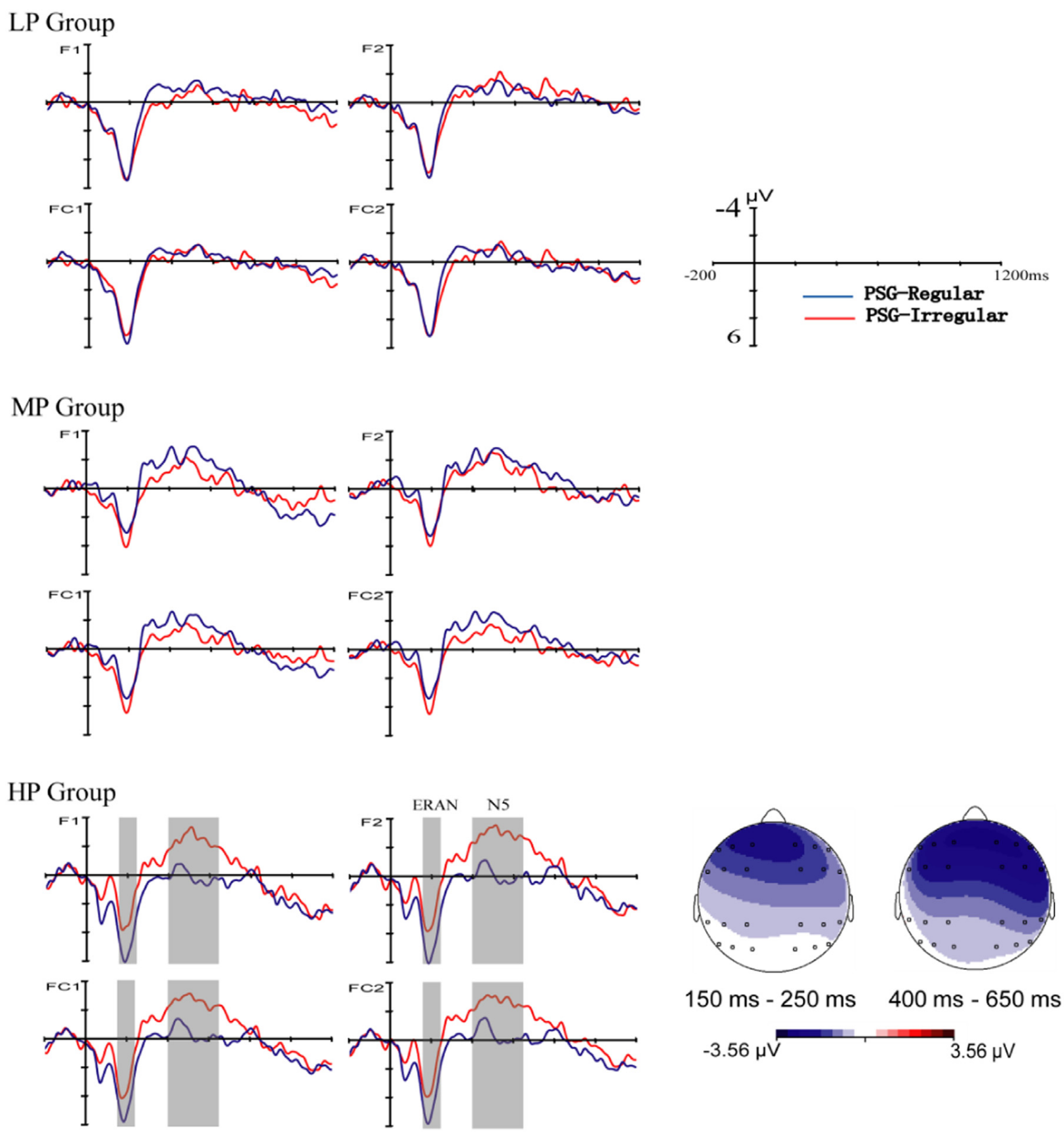

Fig. 4. ERP responses to the final chord in PSG sequences. Event-related potentials (ERPs) evoked by the final chords in PSG sequences are shown separately for low-, moderate-, and high-proficiency groups. For the high-proficiency group, compared to the responses to the regular sequences (blue waveforms), the responses to irregular sequences (red waveforms) evoked an early negativity (maximal around 150–250 ms) and a later negativity (maximal around 400–650 ms). Shaded areas indicate time windows used for the statistical analysis reported in the text. For the low- and moderate-proficiency groups, none of the ERP effects differed significantly between the regular and irregular versions. The right panel shows the scalp distribution of the early and late ERP effects, i.e., the difference potentials between regular and irregular sequences.

anterior negativity (around 150–250 ms after the chord onset) in all three groups; however, a later anterior negativity (around 500–600 ms after the chord onset) was observed only in moderate- and high-proficiency groups (Fig. S2).

3.2.2.1. Results in the 150–250 ms latency range. An analogy global ANOVA for a time window from 150 to 250 ms (early negativity) was carried out (the effects can be seen Fig. S2), indicated a main effect of sequence type [F (1, 46) = 79.28, $p < .001$, $\eta_p^2 = 0.63$] and an interaction among sequence type and anterior–posterior [F (2, 46) = 13.59, $p = .001$, $\eta_p^2 = 0.23$]. Simple effect analyses showed that the negativity was larger for the irregular than for the regular sequence and the effect was mainly distributed in anterior regions ($p < .001$)

3.2.2.2. Results in the 500–600 ms latency range. An analogy global ANOVA for a time window from 500 to 600 ms was carried out (the effects can be seen Fig. S2), indicated a main effect of sequence type [F (1, 46) = 12.47, $p = .001$, $\eta_p^2 = 0.21$] and an interaction among sequence type, anterior–posterior and group [F (2, 46) = 3.69, $p = .032$, $\eta_p^2 = 0.14$]. Simple effect analyses showed that for low-proficiency group, no significant effect of sequence type was found; for moderate-proficiency ($p < .001$) and high-proficiency group ($p < .001$), the negativity was larger for the irregular than for the regular sequence, and the effect was mainly distributed in anterior region.

Table 2
Mean amplitudes (SEM) of ERP effects on final chord.

	150–250 ms Frontal	400–650 ms Frontal
FSG		
LP Group	− 0.001 (0.519)	0.241 (0.513)
MP Group	**− 1.873 (0.519)**	**− 1.912 (0.513)**
HP group	**− 1.112 (0.503)**	**− 2.194 (0.498)**
PSG		
LP Group	0.185 (0.489)	− 0.053 (0.587)
MP Group	0.302 (0.489)	0.573 (0.578)
HP Group	**− 1.508 (0.475)**	**− 2.282 (0.569)**

Mean amplitudes (with SEM) of ERP effects (regular subtracted from irregular events) for the two syntax types (FSG and PSG) and three groups (low-proficiency, moderate-proficiency, and high-proficiency) at the frontal ROIs. Bold font indicates that the amplitude differences between the regular and irregular versions were statistically significant at the frontal ROIs ($p < .05$).

4. Discussion

Both electrophysiological and behavioral data showed that Chinese listeners' Western musical syntax processing was influenced by the structural complexity and Western musical proficiency and that these two factors interacted with each other. In the FSG condition, the event-related potentials (ERP) evoked by regular and irregular final chords were not significantly different in the low-proficiency group. In contrast, in the moderate- and high-proficiency groups, the irregular final chord evoked an ERAN-N5 biphasic response. In the PSG condition, however, only the high-proficiency group showed an ERAN-N5 biphasic response evoked by irregular final chords. These ERP results could gain support from the conclusiveness ratings obtained after the EEG session. The details are discussed in the following sections.

4.1. Effect of structural complexity on the processing of non-native music syntax

First, the findings of the present study revealed that for Chinese listeners, the integration of Western tonal syntax of different structural complexity had different requirements for listener's proficiency and experience. The ERAN-N5 biphasic effect has been taken as an index of the sensitivity to the processing of tonal structure (Koelsch et al., 2000, 2007; Koelsch and Jentschke, 2009). In the FSG condition, no ERP components were elicited by irregular final chords in the low-proficiency group; meanwhile, in the PSG condition, no ERP components were elicited in either the low- or moderate-proficiency groups. The behavioral data also showed that there are no differences in the conclusiveness ratings between irregular and regular sequences for the low-proficiency group in the FSG condition, and for both the low- and moderate-proficiency groups in the PSG condition. That is, the low-proficiency group is insensitive to syntactic violation in the FSG condition; while, both the low- and moderate-proficiency groups were insensitive to syntactic violations in the PSG condition during on-line processing. In the present study, the low-proficiency group included listeners who have no explicit knowledge of Western musical syntax and seldom listen to Western music. The moderate-proficiency group included listeners who have no explicit knowledge of Western musical syntax but frequently listen to Western music. Lastly, the high-proficiency group included listeners who have received many years of formal or professional training in Western music. Therefore, their musical syntax processing performance reflected by different ERP patterns demonstrates that for Chinese listeners, FSG in Western music could be acquired through long-term implicit learning; however, more complex structures, such as PSG, may be difficult to acquire merely through accumulation of the implicit experience.

The results of the present study seem to be inconsistent with findings of the study on native music learners, in which it was reported that for native listeners of Western music the sensitivity to tonal structure can be acquired through implicit learning, both for the FSG (Bigand et al., 1999, 2014; Bigand and Pineau, 1997; Koelsch et al., 2000, 2007) and the PSG (Koelsch et al., 2013; Rohrmeier and Cross, 2009; Woolhouse et al., 2016); in addition, learning the FSG does not require a long-term accumulation of implicit experiences (Jentschke et al., 2014). The discrepancy between our findings on Chinese listeners and the findings on native-music listeners indicates that compared with native listeners, non-native listeners may explore more complicated neural mechanism and have more difficulties in harmonic tonal structure processing and acquisition, especially for complex PSG structures that are not easily acquired merely from implicit learning. Growing evidence from studies on non-native listeners has revealed that cultural bias influences music cognition (Demorest et al., 2016; Patel and Demorest, 2013; Matsunaga et al., 2014). Even when non-native listeners were fully exposed to a non-native music system, they still showed lower musical competence relative to native listeners (Demorest et al., 2016; Demorest and Osterhout, 2012). However, our results extend these findings by showing that the complexity of musical structures would enlarge the gap between the processing of non-native and native music.

Bilingual studies have revealed that for non-native learners syntax is the outstanding challenge among aspects of language learning (Clahsen and Felser, 2006). It has been shown that even some highly proficient learners still have difficulties in on-line non-native syntax processing and find it hard to reach a high degree of automaticity, similar to native speakers (Hahne and Friederici, 2001; Felser and Roberts, 2007; Foucart and Frenck-Mestre, 2012). These may be ascribed to factors such as missing the critical learning period (DeKeyser and Larson-Hall, 2005), inhibiting the interference of the native language (Foucart and Frenck-Mestre, 2011; Kim et al., 2015), and lower exposure rates (Clahsen and Felser, 2006), among others. In this case, the demand on general cognitive resources in syntax processing is generally larger than in one's native language processing. The shortage of cognitive resources may be the reason for non-native learners failing to complete the in-depth structural construction (Clahsen and Felser, 2006; Dowens et al., 2010; Mcdonald, 2006; Suh et al., 2007; Zhou et al., 2017), especially when the structure is more complex and thus more cognitive resources are required (Clahsen and Felser, 2006; Bañón et al., 2012; Hasegawa et al., 2002; Suh et al., 2007; Yokoyama et al., 2006).

In parallel to non-native language learning, the listeners' performances in this study were poorer when the musical syntactic complexity increased. Integration of FSG structure could be completed by moderate- and high-proficiency groups, whereas only the high-proficiency group could integrate PSG. Language studies have revealed that the PSG structure is more complicated than FSG because of two properties. One is the recursive structures; that is, there are subordinated sentences inserted into the main sentence. Another is the longer distance between integrated elements because of the sentence insertion (Makuuchi et al., 2009, 2013). The long-distance dependence may increase the working memory load (Makuuchi et al., 2013). Even in FSG processing, a longer distance between integrated elements could increase memory load because of the need for maintaining words in working memory, impeding the construction of in-depth structural representation (Bañón et al., 2012; Dowens et al., 2010). Furthermore, the integration of the recursive structure causes an increase in computational complexity and greater involvement of other cognitive control functions such as inhibitory control, attention control, and monitoring (Badre, 2008; Jeon and Friederici, 2015; Koechlin and Summerfield, 2007).

It was our speculation that the difficulty in PSG processing for Chinese listeners may also related to the longer integration distance and complicated cognitive computation. The reasons are, first, we found that the low-proficiency group was not unresponsive to adjacent harmonic violation; the irregularity of the first chord of the second phrase evoked a larger ERAN, although it did not evoke a larger N5 (the ERAN-N5 biphasic effect was observed in the moderate- and high-proficiency groups). However, when the integration distance increases and individuals need to deal with the irregular long-distance dependence of syntactic relations (at the final chord), the effect of the ERAN disappeared for the low-proficiency group (even for simple FSG sequences). This demonstrated that, low-proficiency individuals could detect the adjacent irregularity but they could not "carry" this perception of violation to the end of the sequence. The increased integration distance may lead to increased difficulty during harmonic syntax processing. Furthermore, we would also like to test the effect of cognitive computing while separating it from the effect of integration distance. We create non-adjacent dependency in the FSG sequence and manipulated the integration distances between integrated elements to ensure that they were the same; therefore, the cost of working memory may have been similar between FSG and PSG conditions (Makuuchi et al., 2009, 2013). In this case, the processing of PSG was still more difficult than that of FSG, as reflected by the different ERP component patterns in the moderate- and high-proficiency groups. This result indicated that the increased complexity in PSG processing may also be related to the increase in demand for other cognitive control functions (i.e., inhibitory control, attention control), in addition to working memory.

4.2. Effects of proficiency on music processing

It has always been a controversial issue whether the perception of a non–native sound system (music or language) is directly related to an individual's experience or proficiency (Clahsen and Felser, 2006; Demorest et al., 2008, 2010; Tang et al., 2016). With Chinese-Western bimusicals as listeners, our data provide evidence that the perception of non-native syntactic structure was highly affected by music proficiency level. For the low-proficiency group that rarely listened to Western music, no ERP component was elicited by the syntactic violation, even in the FSG condition, indicating that this group has very limited ability to integrate harmonic structure in real time. With the proficiency level improves, native-like neural responses to irregularities appear gradually; ERAN-N5 responses were elicited for the moderate-proficiency group in the FSG condition and for the high-proficiency group in both the FSG and PSG conditions. Moreover, a significant interaction between the complexity and proficiency effects on non-native syntax processing was found, indicating that the integration of more complex structure required higher proficiency levels.

Researchers assume that the effect of proficiency might be related to the increase in the automaticity of second language syntax processing and thus a decrease in the cognitive demand required by in-depth syntactic representation (Coughlin and Tremblay, 2013; Dowens et al., 2010; Jeon and Friederici, 2015; Service et al., 2002). The relationship between the reduction in the cost of cognition and the increase in the proficiency level of non-native syntax processing is evidenced by the observation that with more practice of sentence comprehension, the processing of L2 requires less WM resources (Service et al., 2002) and other cognitive control functions (Jeon and Friederici, 2013, 2015). Similar to the findings in non-native language studies, in this study it was shown that non-native music syntax processing could gradually approach the level of native listeners with improvement in the proficiency level. The competence of the high-proficiency group to integrate PSG could be related to their higher automaticity of non-native music syntax processing, reducing efforts in complicated structure processing and reliance on working memory and cognitive control functions during syntax processing.

Researchers have also argued that the improvement in proficiency with professional training may have promoted learners' syntactic awareness in hierarchical structure processing, enhancing the effective use of syntactic cues and inhibiting interference from irrelevant information (Bouwer et al., 2016; Ellis et al., 2009; Morganshort et al., 2012; Norris and Ortega, 2000). In this study, compare to the moderate and low groups, the high proficiency participants had a much greater extent in receiving explicit instruction of Western musical knowledge, and also in instrument performance and notation literacy trainings. Any one of these (or some combination) could increase non-native music learners' syntactic awareness and facilitate the responses to more complex musical structures. Zhang et al. (2016) conducted a study on Chinese learners' perception of Western tonal music and showed that non-musicians with little implicit learning experience could only perceive larger music boundaries, and only musicians could perceive the hierarchical organization of musical boundaries. It was evident that high proficiency levels with more professional training could improve non-native music learners' capability to deal with complicated hierarchical structures.

4.3. Theoretical implications

Our findings are important for several reasons. First, a few studies have suggested that music perception is "within cultural borders," as opposed to the view of being "across cultures" (Demorest et al., 2016; Patel and Demorest, 2013; Wong et al., 2009), which was confirmed by our study. Chinese listeners have difficulty integrating syntactic structures in Western music, especially for those listeners without formal training. This is significantly different from the processing of native music listeners. From this point of view, non-native music learning is very similar to non-native language learning.

Second, researchers are concerned with which types of music are more difficult for non-native learners. We found that the complexity of the syntactic structure may be a primary factor affecting efficient non-native music processing. However, the improved proficiency levels can promote the processing of complex syntax. From this point, the acquisition of non-native music syntax is also very similar to the acquisition of non-native language syntax. Besides, this research showed the interactive influence of syntactic complexity and proficiency levels on non-native learners' syntactic processing, suggesting that the two effects may coordinate to influence the non-native syntactic processing processes.

As to the neurocognitive mechanism that causes the difficulty in non-native music acquisition and processing among learners, the poorer performance in PSG processing appeared in our study when the working memory load could be similar to the FSG, revealing that dealing with the increase in complexity may need more involvement of other cognitive control functions, in addition to working memory.

5. Conclusion

Taking Chinese listener as an example, this study aimed to investigate how structural complexity and musical proficiency modulate neural processing of syntax in non-native music. The results showed that syntax complexity and proficiency level influence Chinese listeners' music syntax processing and the two factors interact with each other. Although simple structures in Western music, such as the FSG, could be acquired through long-term implicit learning, complex structures, such as the PSG, may not be as easily acquired merely from exposure. This study provides evidence that compared to the acquisition and processing of native music syntax, the acquisition and processing of non-native music syntax, especially for abstract and complex structures, may involve more complicated neural mechanism and pose more difficulties.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant nos. 31470985, 31660282) and Yunnan Philosophy and Social Science Planning Project (QN2017043). The authors are grateful to Dr. Xiao Xi Wang for valuable suggestions on the manuscript and Dr. Peng Li for instruction on the analysis of data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuropsychologia.2018.10.005.

References

Akrami, H., Moghimi, S., 2017. Culture modulates the brain response to harmonic violations: an eeg study on hierarchical syntactic structure in music. Front. Hum. Neurosci. 11, 00591.
Badre, D., 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn. Sci. 12 (5), 193.
Bañón, J.A., Fiorentino, R., Gabriele, A., 2012. The processing of number and gender agreement in spanish: an event-related potential investigation of the effects of structural distance. Brain Res. 1456 (15), 49–63.
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.
Bigand, E., Delbé, C., Poulincharronnat, B., Leman, M., Tillmann, B., 2014. Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. Front. Syst. Neurosci. 8 (94) (94–94).
Bigand, E., Madurell, F., Tillmann, B., Pineau, M., 1999. Effect of global structure and temporal organization on chord processing. J. Exp. Psychol. Hum. Percept. Perform. 25 (25), 184–197.
Bigand, E., Pineau, M., 1997. Global context effects on musical expectancy. Percept. Psychophys. 59 (7), 1098–1107.
Brattico, E., Tupala, T., Glerean, E., Tervaniemi, M., 2013. Modulated neural processing of western harmony in folk musicians. Psychophysiology 50 (7), 653–663.
Clahsen, H., Felser, A.C., 2006. How native-like is non-native language processing? Trends Cogn. Sci. 10 (12), 564.
Collins, T., Tillmann, B., Barrett, F.S., Delbé, C., Janata, P., 2014. A combined model of sensory and cognitive representations underlying tonal expectations in music: from audio signals to behavior. Psychol. Rev. 121 (1), 33–65.
Coughlin, C.E., Tremblay, A., 2013. Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. Appl. Psycholinguist. 34 (3), 615–646.
DeKeyser, R., Larson-Hall, J., 2005. What does the critical period really mean? In: Kroll, J., deGroot, A.M.B. (Eds.), Handbook of Bilingualism. Oxford University Press, Oxford, pp. 88–108.
Demorest, S.M., Morrison, S.J., Nguyen, V.Q., Bodnar, E.N., 2016. The influence of contextual cues on cultural bias in music memory. Music Percept. 33 (5), 590–600.
Demorest, S.M., Morrison, S.J., Jungbluth, D., Beken, M.N., 2008. Lost in translation: an enculturation effect in music memory performance. Music Percept. Interdiscip. J. 25 (3), 213–223.
Demorest, S.M., Morrison, S.J., Stambaugh, L.A., Beken, M., Richards, T.L., Johnson, C., 2010. An fmri investigation of the cultural specificity of music memory. Social. Cogn. Affect. Neurosci. 5 (2–3), 282.
Demorest, S.M., Osterhout, L., 2012. Erp responses to cross-cultural melodic expectancy violations. Ann. N. Y. Acad. Sci. 1252 (1), 152.
Dowens, M.G., Vergara, M., Barber, H.A., Carreiras, M., 2010. Morphosyntactic processing in late second-language learners. J. Cogn. Neurosci. 22 (8), 1870.
Eerola, T., Louhivuori, J., Lebaka, E., 2009. Expectancy in sami yoiks revisited: the role of data-driven and schema-driven knowledge in the formation of melodic expectations. Musica. Sci. 13 (2), 39–70.
Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., Reinders, H., 2009. Implicit and Explicit Knowledge in Second Language Learning, Testing, and Teaching. Multilingual Matters, Bristol, UK.
Fan, Z.Y., 2003. The Theory and Method of Chinese Pentatonic Harmony. Shanghai music publishing house, Shanghai.
Farbood, M.M., Heeger, D.J., Marcus, G., Hasson, U., Lerner, Y., 2015. The neural processing of hierarchical structure in music and speech at different timescales. Front. Neurosci. 9, 157.
Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., Gibson, E., 2009. Structural integration in language and music: evidence for a shared system. Mem. Cogn. 37 (1), 1–9.
Felser, C., Roberts, L., 2007. Processing wh-dependencies in a second language: a cross-modal priming study. Second Lang. Res. 23 (1), 9–36.
Foucart, A., Frenck-Mestre, C., 2011. Grammatical gender processing in L2: electrophysiological evidence of the effect of L1–L2 syntactic similarity. Biling.: Lang. Cogn. 14 (3), 379–399.
Foucart, A., Frenck-Mestre, C., 2012. Can late l2 learners acquire new grammatical features? Evidence from erps and eye-tracking. J. Mem. Lang. 66 (1), 226–248.
Gratton, G., Coles, M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55, 468–484.
Greenhouse, S.W., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 24 (2), 95–112.
Hasegawa, M., et al., 2002. An fMRI study of bilingual sentence comprehension and workload. Neuroimage 15, 647–660.
Hahne, A., Friederici, A.D., 2001. Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. Biling.: Lang. Cogn. 4 (2), 123–141.
Jentschke, S., Friederici, A.D., Koelsch, S., 2014. Neural correlates of music-syntactic processing in two-year old children. Dev. Cogn. Neurosci. 9(C), 200.
Jeon, H.A., Friederici, A.D., 2013. Two principles of organization in the prefrontal cortex are cognitive hierarchy and degree of automaticity. Nat. Commun. 4 (3), 2041.
Jeon, H.A., Friederici, A.D., 2015. Degree of automaticity and the prefrontal cortex. Trends Cogn. Sci. 19 (5), 244.
Keppel, G., 1991. Design and Analysis: A Researcher's Handbook. Prentice-Hall, Inc., New Jersey.
Kim, E., Baek, S., Tremblay, A., 2015. The role of island constraints in second language sentence processing. Lang. Acquis. 22 (4), 384–416.
Koelsch, S., Gunter, T., Friederici, A.D., Ger, E., 2000. Brain indices of music processing: “nonmusicians” are musical. J. Cogn. Neurosci. 12 (3), 520–541.
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.
Koelsch, S., Jentschke, S., 2009. Differences in electric brain responses to melodies and chords. J. Cogn. Neurosci. 22 (10), 2251–2262.
Koelsch, S., Rohrmeier, M., Torrecuso, R., Jentschke, S., 2013. Processing of hierarchical syntactic structure in music. Proc. Natl. Acad. Sci. USA 110 (38), 15443–15448.
Koechlin, E., Summerfield, C., 2007. An information theoretical approach to prefrontal executive function. Trends Cogn. Sci. 11 (6), 229.
Krumhansl, C.L., Toivanen, P., Eerola, T., Toiviainen, P., Jarvinen, T., Louhivuori, J., 2000. Cross-cultural music cognition: cognitive methodology applied to north sami yoiks. Cognition 76 (1), 13–58.
Lerdahl, F., Jackendoff, R.S., 1987. A generative theory of tonal music. Musicol. Aust. 9 (1), 72–73.
Li, Y.H., 2001. Han Chinese Mode and Harmony. Shanghai music publishing house, Shanghai.
Luck, S.J., 2014. An Introduction to the Event-Related Potential Technique. MIT press, Cambridge, MA.
Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A.D., Ungerleider, L.G., 2009. Segregating the core computational faculty of human language from working memory. Proc. Natl. Acad. Sci. USA 106 (20), 8362.
Makuuchi, M., Grodzinsky, Y., Amunts, K., Santi, A., Friederici, A.D., 2013. Processing noncanonical sentences in broca's region: reflections of movement distance and type. Cereb. Cortex 23 (3), 694–702.
Matsunaga, R., Yokosawa, K., Abe, J., 2012. Magnetoencephalography evidence for different brain subregions serving two musical cultures. Neuropsychologia 50 (14), 3218–3227.
Matsunaga, R., Yokosawa, K., Abe, J., 2014. Functional modulations in brain activity for the first and second music: a comparison of high- and low-proficiency bimusicals. Neuropsychologia 54 (1), 1–10.
Mcdermott, J.H., Schultz, A.F., Undurraga, E.A., Godoy, R.A., 2016. Indifference to dissonance in native amazonians reveals cultural variation in music perception. Nature 535 (7613), 547–550.
Mcdonald, J.L., 2006. Beyond the critical period: processing-based explanations for poor grammaticality judgment performance by late second language learners. J. Mem. Lang. 55 (3), 381–401.
Morganshort, K., Steinhauer, K., Sanz, C., Ullman, M.T., 2012. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. J. Cogn. Neurosci. 24 (4), 933–947.
Nan, Y., Knösche, T.R., Zysset, S., Friederici, A.D., 2008. Cross-cultural music phrase processing: an fmri study. Hum. Brain Mapp. 29 (3), 312–328.
Norris, J.M., Ortega, L., 2000. Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. Lang. Learn. 50, 417–528.
Patel, A.D., 2003. Language, music, syntax and the brain. Nat. Neurosci. 6, 674–681. https://doi.org/10.1038/nn1082.
Patel, A.D., Demorest, S.M., 2013. Comparative music cognition: cross-species and cross-cultural studies. In: Deutsch, D. (Ed.), The Psychology of Music, 3rd ed. Academic Press, San Diego, CA, pp. 647–681.
Przysinda, E., Zeng, T., Maves, K., Arkin, C., Loui, P., 2017. Jazz musicians reveal role of expectancy in human creativity. Brain Cogn. 119, 45–53.
Rohrmeier, M., 2011. Towards a generative syntax of tonal harmony. J. Math. Music 5 (1), 35–53.
Rohrmeier, M., Cross, I., 2009. Tacit tonality: implicit learning of context-free harmonic structure. In: Louhivuori, J., et al. (eds.), Proceedings of the 7th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM 2009), pp. 443–452.
Rohrmeier, M., Dienes, Z., Guo, X., Fu, Q., 2014. Implicit Learning and Recursion. Springer, New York.
Rohrmeier, M., Rebuschat, P., 2012. Implicit learning and acquisition of music. Top. Cogn. Sci. 4 (4), 525.
Rossi, S., Gugler, M.F., Friederici, A.D., Hahne, A., 2006. The impact of proficiency on syntactic Second language processing of German and Italian: evidence from event-related potentials. J. Cogn. Neurosci. 18 (12), 2030–2048.
Service, E., Simola, M., Metsänheimo, O., Maury, S., 2002. Bilingual working memory span is affected by language skill. Eur. J. Cogn. Psychol. 14 (3), 383–408.
Slevc, L.R., Okada, B.M., 2015. Processing structure in language and music: a case for shared reliance on cognitive control. Psychon. Bull. Rev. 22 (3), 1–16.
Suh, S., Yoon, H.W., Lee, S., Chung, J.Y., Cho, Z.H., Park, H., 2007. Effects of syntactic complexity in L1 and L2; an fmri study of korean-english bilinguals. Brain Res. 1136 (1), 178–189.
Tang, W., Xiong, W., Zhang, Y.X., Dong, Q., Nan, Y., 2016. Musical experience facilitates lexical tone processing among Mandarin speakers: behavioral and neural evidence. Neuropsychologia 91, 247–253.
Wong, P.C., Chan, A.H., Roy, A., Margulis, E.H., 2011. The bimusical brain is not two monomusical brains in one: evidence from musical affective processing. J. Cogn. Neurosci. 23 (12), 4082–4093.
Wong, P.C., Roy, A.K., Margulis, E.H., 2009. Bimusicalism: the implicit dual enculturation

of cognitive and affective systems. Music Percept. Interdiscip. J. 27 (2), 81–88.
Woolhouse, M., Cross, I., Horton, T.J., 2016. Perception of non-adjacent tonic key relationships. Psychol. Music 44 (4), 1–14.
World Medical Organization, 1996. Declaration of Helsinki (1964). Br. Med. J. 313, 1448–1449.
Xu, F., 2014. Inheritance and Innovation of Chinese Traditional Music Pitch Elements in Contemporary Music. People's Music Publishing House, Beijing.
Yokoyama, S., Okamoto, H., Miyamoto, T., Yoshimoto, K., Kim, J., Iwata, K., et al., 2006. Cortical activation in the processing of passive sentences in L1 and L2: an fmri study. Neuroimage 30 (2), 570–579.
Zhang, J., Jiang, C., Zhou, L., Yang, Y., 2016. Perception of hierarchical boundaries in music and its modulation by expertise. Neuropsychologia 91, 490–498.
Zhang, J., Zhou, X., Chang, R., Yang, Y., 2017. Effects of global and local contexts on chord processing: an erp study. Neuropsychologia 109, 149–154.
Zhou, H., Rossi, S., Chen, B., 2017. Effects of working memory capacity and tasks in processing L2 complex sentence: evidence from Chinese-English bilinguals. Front. Psychol. 8, 1–10.