Musical training modulates the early but not the late stage of rhythmic syntactic processing

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/psyp.12983

Publisher: Wiley-Blackwell

www.reading.ac.uk/centaur

CentAUR

Musical training modulates the early but not the late stage of rhythmic syntactic processing

Lijun Sun[a], Fang Liu[b], Linshu Zhou[c], Cunmei Jiang[c*]

[a] College of Education, Shanghai Normal University, Shanghai, China

[b] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

[c] Music College, Shanghai Normal University, Shanghai, China

Running head: Music training and rhythmic syntactic processing

[*] Corresponding author at: Music College, Shanghai Normal University, 100 Guilin Road, Xuhui District, Shanghai 200234, China.

E-mail: cunmeijiang@126.com (C. Jiang)

Abstract

Syntactic processing is essential for musical understanding. Although the processing of harmonic syntax has been well studied, very little is known about the neural mechanisms underlying rhythmic syntactic processing. The present study investigated the neural processing of rhythmic syntax and whether and to what extent long-term musical training impacts such processing. Fourteen musicians and 14 nonmusicians listened to syntactic-regular or -irregular rhythmic sequences and judged the completeness of these sequences. Musicians, as well as nonmusicians, showed a P600 effect to syntactic-irregular endings, indicating that musical exposure and perceptual learning of music are sufficient to enable nonmusicians to process rhythmic syntax at the late stage. However, musicians, but not nonmusicians, also exhibited an ERAN response to syntactic-irregular endings, which suggests that musical training only modulates the early but not the late stage of rhythmic syntactic processing. These findings revealed for the first time the neural mechanisms underlying the processing of rhythmic syntax in music, which has important implications for theories of hierarchically-organized music cognition and comparative studies of syntactic processing in music and language.

Keywords: rhythmic syntax, musical training, cognition, ERAN, P600

1. Introduction

As music unfolds in time, temporal processing becomes a critical aspect in music cognition. Sensitivity to the basic temporal information of music is universal across human and nonhuman species (Cook, Rouse, Wilson, & Reichmuth, 2013; Patel, 1998; Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Research has shown that sulphur-crested cockatoos (Patel, Iversen, Bregman, & Schulz, 2009) and California sea lions (Cook et al., 2013) can synchronize movements to music (see also Merker, Morley, & Zuidema, 2015 for a review), while newborn infants are able to detect the beat in a sequence (Winkler et al., 2009). It has also been demonstrated that adult listeners can discriminate the difference between rhythmic sequences, either in the metric simple condition, in which accents occur at regular intervals (Grahn & Brett, 2007; Grahn & Schuit, 2012), or in the metric complex condition, in which accents occur at irregular intervals (Grahn & Brett, 2007; Hopyan, Schellenberg, & Dennis, 2009). Although both musicians and nonmusicians show sensitivity to violation of rhythm and beat by eliciting a mismatch negativity (MMN) component (Rüsseler, Altenmüller, Nager, Kohlmetz, & Münte, 2001; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009; Vuust et al., 2005), musical training can facilitate rhythm and beat perception (Geiser, Ziegler, Jancke, & Meyer, 2009; Habibi, Wirantana, & Starr, 2014), with musicians showing a larger MMN amplitude than nonmusicians (Rüsseler et al., 2001; Vuust et al., 2009; Vuust et al., 2005). The MMN is an auditory event-related potential elicited by infrequent auditory stimuli deviating from a regular standard sound in terms of a specific physical feature, reflecting early, pre-attentive sound discrimination abilities (Kujala, 2007).

Rhythm is organized around an underlying metrical structural framework consisting of

strong (accented) and weak (unaccented) beats (Handel, 1989). Thus, rhythm is processed in parallel with beat in music perception. In an oddball detection task, an MMN component was elicited in normal listeners when a metrically weak position was placed with an accented event (Geiser, Sandmann, Jäncke, & Meyer, 2010; Vuust et al., 2009; Vuust et al., 2005), and when a rhythmic event in a metrically strong position was omitted (Bouwer, Van Zuijen, & Honing, 2014; Honing, Ladinig, Háden, & Winkler, 2009; Ladinig, Honing, Háden, & Winkler, 2009; Winkler et al., 2009). Compared to nonmusicians, musicians showed a larger MMN amplitude in the oddball tasks, suggesting greater sensitivity to the mismatch between an accented event and a metrically weak position in musicians than nonmusicians (Geiser et al., 2010; Vuust et al., 2009; Vuust et al., 2005). In this case, the MMN component represented the perception of deviant rhythms and reflected sensory information processing, as it was elicited by sensory dissonance and the low probability of the deviant rhythms.

Given that metrically strong/weak positions are constant in a piece of music, a rhythmic accent occurring at a metrically weak position would result in a mismatch. In music, syncopation refers to the rhythmic pattern in which accented/strong events are placed at metrically weak positions, or no/weak events are placed at strong beat locations in the underlying metrical structure (Fitch & Rosenfeld, 2007). Syncopation occurring in the middle of a phrase would lead to sensory dissonance. However, when it occurs at the end of a phrase, a violation of rhythmic syntax would arise. This is because the syncopation results in the absence of a rhythmic accent in the last metrically strong position of a melody, given that ending a melody with an accented event at the last metrically strong position is required by regular rhythmic syntax.

Apart from harmonic syntax, rhythmic syntax is also an important form of musical syntax (Fitch, 2013; Koelsch, 2013). Similar to harmonic syntax, rhythmic syntax reflects the hierarchy of the temporal structure in music, in which discrete elements are organized in a

4

subordinate or dominant way based on the relationship between strong and weak events (Fitch, 2013). A rhythmic accent placed in a metrically strong position at the end of a phrase sounds more complete (and thus regular), and makes listeners feel stable, while a rhythmic accent placed in a metrically weak position at the end of a phrase sounds incomplete (and thus irregular), and makes listeners feel unstable. This suggests that rhythmic syntactic processing relies on the representations of regularities in long-term memory, which is at the cognitive processing stage and beyond the sensory mode.

Although musical syntax comprises regularities in melody, rhythm, harmony, and timbral structure, etc. (Koelsch, 2013), previous studies mainly focused on harmonic syntax in tonal music (e.g., Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch & Jentschke, 2008, 2010; Steinbeis, Koelsch, & Sloboda, 2006), and showed that syntactically irregular chords and tones elicited an early right anterior negativity (ERAN) and P600. The ERAN has an anterior right-hemispheric scalp distribution with a latency from 100 to 350 msec, which reflects early automatic neural processing of musical syntax (Koelsch et al., 2000; Koelsch & Jentschke, 2010; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Steinbeis et al., 2006). The P600 appears mostly on the posterior part of the center of the scalp with an onset around 300 msec after the irregular chords or tones and reflects late conscious neural processing (Besson & Faïta, 1995; Patel, 1998; Regnault, Bigand, & Besson, 2001; Zendel, Lagrois, Robitaille, & Peretz, 2015). Unlike harmonic syntax, however, very little is known about the neural substrates of rhythmic syntactic processing.

The present study used event-related potentials to investigate the neural substrates of rhythmic syntactic processing and the extent to which long-term musical training affects such processing. Syntactically regular rhythmic sequences were composed to have a rhythmic accent in a metrically strong position at the end of a phrase. In contrast, using syncopation, syntactically irregular sequences ended with a rhythmic accent in the last metrically weak

position of the penultimate bar, which lasted until the first metrically strong position of the final bar. This would allow us to compare brain responses to the final tone of a sequence on a strong beat and a sustained note from the last note of the penultimate bar on a strong beat. We hypothesized that the violation of rhythmic syntactic regularity would elicit a P600 in both musicians and nonmusicians, which is a neurophysiological marker of chord-syntactic processing. We also predicted that musicians would show a different ERAN effect from nonmusicians, given that musicians have been found to have larger and earlier brain responses to harmonic syntax than nonmusicians (Koelsch, Jentschke, Sammler, & Mietchen, 2007; Koelsch, Schmidt, & Kansok, 2002).

2. Method

2.1 Participants

Fourteen musicians (M_{age} = 23.21 years, SD = 2.49, 11 females) and 14 nonmusicians (M_{age} = 23.86 years, SD = 2.51, 11 females) participated in the experiment. Nonmusicians did not receive any extracurricular training in music. Musicians had received, on average, 16 years of formal instrumental training (*range: 8 to 25*), and played either piano, violin, viola, accordion, cello, tuba, erhu (a traditional Chinese instrument), or Chinese flute. All participants were right-handed and reported to have 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 Stimuli

Three original regular rhythmic sequences were composed with a fixed pitch (A2) in 2/4

meter. Each sequence consisted of eight bars. Only simple notes, such as quarter note, eighth note, sixteenth note and half note, were included. The duration of each sequence was about 11 seconds. Each sequence was given a regular and an irregular ending. Sequences with regular endings had the terminal event in the first metrically strong position of the final bar (see Figure 1A for an example). The endings with congruous rhythmic and metric accents represented a stable completion. Sequences with irregular endings had the terminal event syncopated, which began in the last metrically weak position of the penultimate bar and lasted until the first metrically strong position of the final bar (see Figure 1B for an example). The endings with incongruous rhythmic and metric accents created an unexpected incompletion. As shown in Figure 1, we compared how the brain responded to the final tone of a sequence on a strong beat and to a sustained note from the last note of the penultimate bar on a strong beat. Because the sustained notes occurred in syntactically irregular sequences, this design would rule out the possibilities that the ERP effects of these irregular events were due to acoustic processing. In order to exclude the possible effects of sensory dissonance on the ERP results, we created syncopation in two other locations (in the second and sixth bars) before the terminal event in each sequence.

INSERT FIGURE1 ABOUT HERE

The six sequences (three original rhythmic sequences with regular and irregular endings) were then transposed to four different pitches in D^b2, E^b2, F2, and G2, yielding 30 sequences. The 30 sequences were repeated three times, resulting in 90 trials in the experiment. All stimuli were created at a tempo of 108 beats per minute using the Sibelius 7.5 software (Avid Tech. Inc.). The conga timbre of the stimuli was generated by Cubase 5.1's inbuilt Kontact 5.4. To ensure that the velocity was constant across all sound events (including the downbeat),

the loudness of each sequence was normalized to -3 dB using Adobe Audition 3.0. In order to help participants form a metrical structural framework, four beats were presented before each sequence in the experimental trials.

2.3 Procedure

All stimuli were presented in a pseudorandom order with the constraints that a given ending (regular or irregular) was not repeated more than three times in succession, and that consecutive sequences were not in the same pitch. The experiment was conducted in an acoustically and electrically shielded room. All stimuli were presented binaurally through Philips SHM1900 headphones. In order to avoid misunderstanding, participants were informed of the meaning of "completion" before testing. During the experiment, participants were required to judge whether the completion of each sequence was fine by pressing one of the two response buttons (yes/no) using their left/right index finger after the presentation of each sequence. The order of 'yes' and 'no' response buttons was counterbalanced across participants. Three practice trials were given before the experimental session to familiarize the participants with the stimuli and procedure.

2.4 ERP Recording and Analysis

EEG data were recorded from 64 standard scalp locations (International 10–20 system), digitized at a rate of 500 Hz, with a 0.05 Hz low cutoff filter and a 100 Hz high cutoff filter. Electrodes were also placed above and below the left eye and on the outer canthi of both eyes to measure vertical and horizontal eye movements and eye-blinks. The data were referenced offline to the algebraical mean of left and right mastoid electrodes, filtered offline with a band-pass filter of 0.1 to 30 Hz (24-dB/oct slope). ERPs time-locked to the onset of the final bar were averaged for each participant under each condition. Epochs of 1200 msec (including

a 200 msec pre-stimulus period) were averaged using a 200 msec pre-stimulus baseline. Trials were excluded from analysis when additional artifacts exceeded the amplitude of ±75 μV in any channel.

ERPs were analyzed statistically in four regions of interest (ROIs): left anterior electrodes (F7, F5, F3, FT7, FC5, FC3), right anterior electrodes (F8, F6, F4, FT8, FC6, FC4), left posterior electrodes (P7, P5, P3, TP7, CP5, CP3), and right posterior electrodes (P8, P6, P4, TP8, CP6, CP4). Based on visual inspection and previous research on the ERAN (Koelsch & Jentschke, 2010; Paller, Mccarthy, & Wood, 1992) and P600 (Besson & Faïta, 1995; Regnault et al., 2001; Zendel & Alain, 2009), a time window from 100 to 160 msec was selected for the analysis of the ERAN, and a time window from 300 to 700 msec was selected for the analysis of the P600. Repeated measures ANOVAs taking regularity (regular, irregular), hemisphere (left, right), anteriority (anterior, posterior) as within-subjects factors and group (musicians, nonmusicians) as the between-subjects factor were conducted. The mean of the respective electrodes in each region of interest was computed for analysis. Only the significant effects containing the main experimental variables (regularity and group) were reported. Overall ANOVAs were followed up by simple effects tests if there were significant two- and three-way interactions between the main experimental variables. It is worth noting that we found no outliers with ERP amplitudes 2 standard deviations above/below the mean value of each group, whether for the time window of 100 to 160 msec, or 300 to 700 msec. Thus, all data were included in statistical analysis.

3. Results

3.1 Behavioral results

Percentages of correct judgments of the completeness of the sequences were calculated for all participants. Both musicians and nonmusicians performed well above chance (musicians: $p < .001$; nonmusicians: $p = .005$). A two-way ANOVA taking group (musicians versus nonmusicians) as the between-subjects factor and regularity (regular versus irregular) as the within-subjects factor was performed. There was a significant main effect of group ($F_{(1, 26)} = 16.58$; $p < .001$, partial $\eta^2 = .39$), reflecting that musicians ($M = 88.25\%$, $SD = 11.08$) outperformed nonmusicians ($M = 65.66\%$, $SD = 17.64$) on completeness judgment. Moreover, there was a main effect of regularity ($F_{(1, 26)} = 7.77$; $p = .01$, partial $\eta^2 = .23$), with regular endings better detected than irregular endings. The interaction was not significant ($p = .16$).

3.2 EEG results

Figures 2 and 3 show the brain electrical responses to syntactic regularity and scalp distribution of the irregular-minus-regular difference waves, respectively. In the time window of 100-160 msec, the syntactic-irregular condition elicited a larger negativity compared to the syntactic-regular condition in musicians, but not in nonmusicians. In the time window of 300-700 msec, the syntactic-irregular condition elicited a larger positivity compared to the syntactic-regular condition in both musicians and nonmusicians (Figures 2-3).

INSERT FIGURE2 ABOUT HERE

For the time window of 100 to 160 msec, a three-way ANOVA revealed an effect of regularity ($F_{(1, 26)} = 11.91$; $p = .002$, partial $\eta^2 = .31$), indicating that syntactic-irregular endings elicited a larger negativity than syntactic-regular endings, and an interaction between regularity and hemisphere ($F_{(1, 26)} = 21.70$, $p < .001$, partial $\eta^2 = .46$), reflecting a right-hemisphere predominance of the negativity elicited by syntactic-irregular endings: $F_{(1, 26)} = 17.83$, $p < .001$, partial $\eta^2 = .41$ for the right hemisphere, $F_{(1, 26)} = 5.41$, $p = .028$, partial $\eta^2 = .17$ for the left hemisphere. Furthermore, there was a three-way interaction among group, regularity, and hemisphere ($F_{(1, 26)} = 11.55$, $p = .002$, partial $\eta^2 = .31$), reflecting that musicians responded to irregular and regular endings differently in both the left hemisphere ($F_{(1, 26)} = 6.02$, $p = .021$, partial $\eta^2 = .19$) and the right hemisphere ($F_{(1, 26)} = 23.30$, $p < .001$, partial $\eta^2 = .47$), whereas nonmusicians did not show different responses to irregular and regular endings for either hemisphere, $ps > .26$.

In order to examine whether the above observed ERP effects in musicians were due to syncopation, rather than to irregular rhythmic syntax at the end of the sequence, we compared brain responses to the sustained notes in the second, sixth, and eighth (the final) bars of the irregular sequences in musicians, by conducting a repeated measures ANOVA taking position (second, sixth, and eighth), hemisphere (left, right) and anteriority (anterior, posterior) as within-subjects factors. As shown in Figure 4, the negativity was the largest for the sustained notes in the eighth bar, followed by the second and then the sixth bar. The ANOVA results showed a two-way interaction between hemisphere and position ($F_{(2, 26)} = 8.18$, $p = .003$, partial $\eta^2 = .39$), reflecting that musicians responded to the three sustained notes differently only in the right hemisphere ($F_{(2, 12)} = 8.58$, $p = .005$, partial $\eta^2 = .59$), but not in the left hemisphere ($F_{(2, 12)} = .28$, $p = .761$, partial $\eta^2 = .05$). In the right hemisphere, the sustained

notes in the final bar elicited a larger negativity than those in the second ($p = .023$) and sixth bars ($p = .003$), while no significant difference was found between the sustained notes in the second and sixth bars ($p = 1.000$).

INSERT FIGURE4 ABOUT HERE

In order to examine whether the absence of the ERAN effect in nonmusicians resulted from the failure to detect syntactic violations, we conducted a repeated measures ANOVA taking detection (detected, undetected), hemisphere (left, right) and anteriority (anterior, posterior) as within-subjects factors. The results showed no significant difference between detected and undetected irregular endings ($p = .47$), indicating that the absence of the ERAN in nonmusicians was not due to their failure to detect irregular endings.

For the time window of 300 to 700 msec, the ANOVA revealed an effect of regularity ($F_{(1,26)} = 14.63$, $p = .001$, partial $\eta^2 = .36$), indicating that syntactic-irregular endings elicited a larger positivity than syntactic-regular endings in both musicians and nonmusicians. Moreover, there was an interaction between regularity and hemisphere ($F_{(1, 26)} = 4.27$, $p = .049$, partial $\eta^2 = .14$), reflecting that syntactic-irregular endings elicited a larger positivity in the left hemisphere ($F_{(1, 26)} = 22.68$, $p < .001$, partial $\eta^2 = .47$) than the right hemisphere ($F_{(1, 26)} = 4.64$, $p = .041$, partial $\eta^2 = .15$). No other significant main effect or interaction was found (all ps $> .070$).

4. Discussion

Although musical syntax comprises harmonic and rhythmic components, little is known about the neural substrates of rhythmic syntax, as compared to harmonic syntax. This study

investigated the neural substrates underlying the processing of rhythmic syntax and whether and to what extent long-term musical training impacts such processing. While both musicians and nonmusicians exhibited a positive component with a latency window of 300-700 msec in response to syntactic-irregular endings, an early right-lateralized negativity in a latency window of 100-160 msec in response to syntactic-irregular endings was observed only in musicians, but not in nonmusicians. This indicates that musical training modulates the processing of rhythmic syntax at the early automatic, but not at the late controlled stage. To the best of our knowledge, this is the first study to examine the neural mechanisms underlying rhythmic syntactic processing and the effect of musical training on such processing.

4.1 The universal processing of rhythmic syntax at the late controlled stage

One of the main findings of our study is that, like musicians, nonmusicians are also able to process rhythmic syntax consciously by showing a late positive component in response to syntactic-irregular endings. This late positive component is reminiscent of the P600 elicited by incongruous chords (Patel, 1998; Regnault et al., 2001) or notes (Besson & Faïta, 1995; Peretz, Brattico, Järvenpää, & Tervaniemi, 2009; Zendel et al., 2015), reflecting conscious processes of structural integration and re-parsing of incongruous events into musical context (Besson & Faïta, 1995; Patel, 1998).

The left-lateralized P600 was, however, different from the centro-parietal distribution of pitch processing in music. This may be attributed to the specificity of temporal processing. Indeed, contrary to pitch processing, brain areas of the left hemisphere involve rhythm processing in recognition (Piccirilli, Sciarma, & Luzzi, 2000), identification (Platel et al., 1997), production (Murayama, Kashiwagi, Kashiwagi, & Mimura, 2004; Piccirilli et al., 2000), and reading tasks (Midorikawa, Kawamura, & Kezuka, 2003). Furthermore,

neuroimaging studies of language processing also suggest the roles of the left hemisphere such as the left temporal gyrus and left basal ganglia in the processing of syntactic violation (Friederici, Rüschemeyer, Hahne, & Fiebach, 2003), particularly during late syntactic revision and integration (Friederici & Kotz, 2003). This late syntactic processing has been indexed as the ERP component P600, reflecting controlled processes of syntactic reanalysis and repair in the language domain (Friederici, Hahne, & Mecklinger, 1996; Hahne & Friederici, 1999). Thus, the P600 effect observed in the present study demonstrates that musicians as well as nonmusicians are able to process rhythmic syntax consciously by re-parsing incongruous rhythmic events into musical context. This finding indicates that musical exposure and perceptual learning of music may be sufficient for the development of more controlled rhythmic syntactic processing mechanisms, and provides evidence for the universal processing of rhythmic syntax at the late controlled stage.

4.2 The effect of musical training on the processing of rhythmic syntax at the early automatic stage

Another significant finding of the present study is that musicians, but not nonmusicians, exhibited an early right-lateralized negativity in a latency window of 100-160 msec in response to syntactic-irregular endings. The latency, morphology, and scalp distribution of this early negativity is highly consistent with the early right anterior negativity (ERAN) elicited by syntactically inappropriate harmonic chords or tones in previous studies (Koelsch & Jentschke, 2008, 2010; Paller et al., 1992), although the ERAN in the present study had a relatively earlier latency of 100 to 160 msec. This pattern might result from the use of repetitive or monophonic music sequences (Koelsch & Jentschke, 2008, 2010).

It is worth noting that the ERAN effect observed in the present study was due to rhythmic-syntactic processing, rather than to the perception of syncopated rhythms. In order

to control for the effect of syncopation, we included three syncopated rhythms in the second, sixth, and final bars of each syntactic-irregular rhythmic sequence. A previous study on chord processing (Koelsch et al., 2000) showed that if there were no syntactic processing, the ERP amplitude elicited by the same chord events would decrease systematically as a function of the positions of the events in the progressing sequences. If this were the case for our current study, the ERP amplitude elicited by the final sustained notes would be the smallest among the three positions. However, our results demonstrated that the final sustained notes elicited the largest ERP amplitude among the three positions. Therefore, this ERP component cannot be simply due to the effect of syncopated rhythm, but to the processing of rhythmic syntax.

Furthermore, this ERAN component could not have simply been an N1 effect, although the two components occur at roughly similar time windows. As shown in Figure 1, the final notes of the regular sequences were on a strong beat, while the sustained notes from the last notes of the penultimate bars were also on a strong beat in the irregular sequences. In this case, the N1 component could not have been evoked by the sustained notes in the irregular sequences.

In addition, this ERAN component was different from the MMN component as exhibited in previous studies (Bouwer et al., 2014; Geiser et al., 2009; Ladinig et al., 2009; Vuust et al., 2009). Two reasons may account for this difference. As noted earlier, the MMN component is elicited by sensory novelty of low probability events (usually 15-25%) in auditory patterns. In the present experiment, however, the frequency of occurrence of rhythmic sequences with syntactic-irregular endings reached 50%, which was the same for rhythmic sequences with syntactic-regular endings. In this case, the ERAN elicited would be due to the novelty of syntactic-irregular endings, but not due to sensory novelty. On the other hand, for the irregular condition, we used a sustained note from the last note of the penultimate bar in the metrically strong position at the end of a sequence, and thus created a

syncopation. The syncopation also occurred in two other locations in each sequence (see

Figure 1). As such, the overall probability for the occurrence of syncopation was above 30%,

which would be too high to elicit an MMN component, since in auditory oddball paradigms,

deviants occurring with a probability of more than 20–25% hardly elicit any deviance-related

negativity (Koelsch et al., 2007; Schröger, 1998). Therefore, unlike the MMN, the ERAN

observed in the present study reflected cognitive processing of complex musical regularities

that existed in long-term memory, rather than due to sensory novelty.

Our findings suggest that musicians, but not nonmusicians, exhibited the ERAN effect in

response to syntactic-irregular endings, which may be attributed to enhanced representations

of rhythmic-syntactic regularities in musicians. The ERAN effect reflects fast and automatic

neural processing related to complex music-syntactic irregularities (Koelsch et al., 2002).

Thus, our results revealed that musical training modulates the early stage of rhythmic

syntactic processing. The effect of musical training is consistent with previous studies

suggesting that musical training has a positive effect on rhythm and beat perception (e.g.,

Geiser et al., 2010; Geiser et al., 2009; Vuust et al., 2009). These findings indicate that

long-term musical training facilitates the processing of temporal information at both

perceptual and cognitive levels. Furthermore, given that a larger ERAN effect has also been

observed in musicians relative to nonmusicians in the processing of harmonic syntax

(Koelsch et al., 2007; Koelsch et al., 2002), the present findings suggest that long-term

musical training affects the processing of music structure in both rhythmic and harmonic

dimensions, at least at the early stage.

Although the elicitation of the ERAN effect was modulated by musical training, the

amplitude of the P600 component did not differ between musicians and nonmusicians. The

differential influences of musical training on the ERAN and P600 effects might be due to the

independence of automatic and controlled syntactic processing. This dissociation between

automatic and controlled syntactic processing has been exhibited in language studies. For example, while both the early left anterior negativity (ELAN) and P600 were elicited in native speakers to syntactic violations in language, only the P600, but not the ELAN, was elicited in nonnative speakers (Hahne, 2001; Mueller, Hahne, Fujii, & Friederici, 2005). The results were explained by the fact that although a more controlled repair integration (reflected by the P600) was accessible to second-language learners, they could not achieve automatic processing (as reflected by the ELAN), indicating that the ELAN effect is not a prerequisite for the elicitation of a P600 effect. This is the case for the ERAN and P600 effects in music. Furthermore, although showing late conscious processing of rhythmic syntax, nonmusicians did not exhibit an early right anterior negativity. The absence of this negativity in nonmusicians was due to their failure to automatically process rhythmic syntax, rather than due to the difficulty in detecting irregular endings.

4.3 Implications for the processing of hierarchical structures in music and language

Musical syntax comprises harmonic and rhythmic syntaxes. Previous studies have reported that the ERAN and P600 effects reflect the processing of harmonic syntax (Besson & Faïta, 1995; Patel, 1998; Peretz et al., 2009; Zendel et al., 2015). Our findings suggest that the ERAN and P600 components can also be regarded as the neurophysiological indexes of rhythmic syntax processing, suggesting that rhythmic syntax is processed in a manner similar to harmonic syntax. Furthermore, analogous to the recursion and hierarchical organization in the harmonic domain (Lerdahl & Jackendo, 1983; Rohrmeier, 2011), the present data support the hypothesis that temporal structure is not simply a flat serial pattern, but rather a hierarchical and recursive organization (Fitch, 2013; Longuet-Higgins, 1979). Together, these findings suggest that musical structure is processed hierarchically in terms of the degree of

stability, both for the harmonic and rhythmic dimensions, providing evidence for the hypothesis of hierarchically-organized cognition in music (Lerdahl & Jackendo, 1983; Rohrmeier, 2011).

On the other hand, the ERAN and P600 elicited by rhythmic syntactic-irregular endings in the present study are reminiscent of the biphasic ERP pattern (ELAN-P600) in the language domain (Friederici & Mecklinger, 1996; Hahne & Friederici, 1999). Indeed, the ELAN and P600 are syntax-related ERP components, signifying automatized structure-building process and late reanalysis and repair in language (Friederici et al., 1996; Hahne & Friederici, 1999). It has been suggested that metric violations elicit an early negativity and a late positivity ERP response (Marie, Magne, & Besson, 2011; Rothermich, Schmidt-Kassow, & Kotz, 2012; Schmidt-Kassow & Kotz, 2009), representing two stages in the processing of metric structure in language. These similarities between music and language may result from more general and abstract principles of hierarchical cognition that apply across domains (Fitch, 2013). Given the interaction between syntactic processing in language and music (Koelsch, Gunter, Wittfoth, & Sammler, 2005), the present findings provide further evidence for the shared neural resources in the processing of syntax in music and language.

5. Conclusion

The present findings revealed for the first time that musical training modulated the early automatic neural processing of rhythmic syntax, although both musicians and nonmusicians were able to process rhythmic syntax consciously. This indicates that musical exposure to and perceptual learning of music are sufficient for the development of controlled rhythmic syntactic processing mechanisms in nonmusicians. However, long-term musical training is

needed for the early automatic and fast processing of rhythmic syntax. Given that the ERAN and P600 are neurophysiological indexes of harmonic syntactic processing as revealed in previous studies (e. g., Koelsch et al., 2000; Koelsch and Jentschke, 2010; Steinbeis et al., 2006), the present study extends this finding to rhythmic syntactic processing, and suggests that musical structure is processed hierarchically in terms of the degree of stability, whether for the harmonic or for the rhythmic dimension. Given the co-occurrence of harmony and rhythm in music, future studies are needed to investigate how harmony interacts with rhythm in the cognition of musical structure.

References

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of
  musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and
  Performance*, *21*(6), 1278–1296. https://doi.org/10.1037/0096-1523.21.6.1278

Bouwer, F. L., Van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for metrically simple
  rhythms with clear accents: An ERP study. *PLoS ONE*, *9*(5), e97467.
  https://doi.org/10.1371/journal.pone.0097467

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can
  keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of
  Comparative Psychology*, *127*(4), 412–427. https://doi.org/10.1037/a0032345

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

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception:
  An Interdisciplinary Journal*, *25*(1), 43–58. https://doi.org/10.1525/MP.2007.25.1.43

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). The temporal structure of syntactic parsing: Early and
  late effects elicited by syntactic anomalies. *Journal of Experimental Psychology: Learning, Memory,
  and Cognition*, *22*(5), 1219–1248. https://doi.org/10.1037/0278-7393.22.5.1219

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion
  studies. *NeuroImage*, *20*(3), S8–S17. https://doi.org/10.1016/S1053-8119(03)00522-6

Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and
  second-pass parsing processes. *Journal of Psycholinguistic Research*, *25*(1), 157–176.
  https://doi.org/10.1007/BF01708424

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and
  superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes.
  Cerebral Cortex, *13*(2), 170–177. https://doi.org/10.1093/cercor/13.2.170

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception-training increases
  hierarchical metre processing. *The European Journal of Neuroscience*, *32*(11), 1979–1985.
  https://doi.org/10.1111/j.1460-9568.2010.07462.x

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and

rhythm processing in music perception. *Cortex*, *45*(1), 93–102.

https://doi.org/10.1016/j.cortex.2007.09.010

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive

Neuroscience*, *19*(5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Grahn, J. A., & Schuit, D. (2012). Individual differences in rhythmic ability: Behavioral and neuroimaging

investigations. *Psychomusicology: Music, Mind and Brain*, *22*(2), 105–121.

https://doi.org/10.1037/a0031188

Habibi, A., Wirantana, V., & Starr, A. (2014). Cortical activity during perception of musical rhythm: Comparing

Musicians and Non-musicians. *Psychomusicology: Music, Mind, and Brain*, *24*(2), 125–135.

https://doi.org/10.1037/pmu0000046

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain

potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266.

https://doi.org/10.1023/A:1010490917575

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early

automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205.

https://doi.org/10.1162/089892999563328

Handel, S. (1989). *Listening*. Cambridge, MA: MIT Press.

Honing, H., Ladinig, O., Háden, G. P., & Winkler, I. (2009). Is beat induction innate or learned? Probing

emergent meter perception in adults and newborns using event-related brain potentials (ERP). *Annals

of the New York Academy of Sciences*, *1169*(1), 93–96.

https://doi.org/10.1111/j.1749-6632.2009.04761.x

Hopyan, T., Schellenberg, E. G., & Dennis, M. (2009). Perception of strong-meter and weak-meter rhythms in

children with spina bifida meningomyelocele. *Journal of the International Neuropsychological Society*,

15(4), 521–528. https://doi.org/10.1017/S1355617709090845

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.

https://doi.org/10.1162/089892900562183

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in

language and music: An ERP-study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–1577.

    https://doi.org/10.1162/089892905774597290

Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain*

    Research, *1212*, 55–62. https://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. https://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.

    https://doi.org/10.1111/j.1469-8986.2007.00517.x

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. https://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.

    https://doi.org/10.1017.S0048577202010508

Kujala, T. (2007). The role of early auditory discrimination deficits in language disorders. *Journal of*

    Psychophysiology, *21*(3-4), 239–250. https://doi.org/10.1027/0269-8803.21.34.239

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in

    adult listeners without extensive music training. *Music Perception*, *26*(4), 377–386.

    https://doi.org/10.1525/MP.2009.26.4.377

Lerdahl, F., & Jackendo, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Longuet-Higgins, H. C. (1979). The perception of music. *Proceedings of the Royal Society of London, Series B*,

    205(1160), 307–322. https://doi.org/10.1098/rspb.1979.0067

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive*

    Neuroscience, *23*(2), 294–305. https://doi.org/10.1162/jocn.2010.21413

Merker, B., Morley, I., & Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music.

    Philosophical Transactions of the Royal Society B, *370*(1664), 20140095.

    https://doi.org/10.1098/rstb.2014.0095

Midorikawa, A., Kawamura, M., & Kezuka, M. (2003). Musical alexia for rhythm notation: A discrepancy

    between pitch and rhythm. *Neurocase*, *9*(3), 232–238. https://doi.org/10.1076/neur.9.3.232.15558

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, *17*(8), 1229–1244. https://doi.org/10.1162/0898929055002463

Murayama, J., Kashiwagi, T., Kashiwagi, A., & Mimura, M. (2004). Impaired pitch production and preserved rhythm production in a right brain-damaged patient with amusia. *Brain and Cognition*, *56*(1), 36–42. https://doi.org/10.1016/j.bandc.2004.05.004

Paller, K. A., Mccarthy, G., & Wood, C. C. (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology*, *29*(2), 202–206. https://doi.org/10.1111/j.1469-8986.1992.tb01686.x

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

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*(10), 827–830. https://doi.org/10.1016/j.cub.2009.05.023

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain*, *132*(5), 1277–1286. https://doi.org/10.1093/brain/awp055

Piccirilli, M., Sciarma, T., & Luzzi, S. (2000). Modularity of music: Evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *69*(4), 541–545. https://doi.org/10.1136/jnnp.69.4.541

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., . . . Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain*, *120*(2), 229–243. https://doi.org/10.1093/brain/120.2.229

Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T. F. (2001). Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neuroscience Letters*, *308*(1), 33–36. https://doi.org/10.1016/S0304-3940(01)01977-2

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255. https://doi.org/10.1162/089892901564298

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53. https://doi.org/10.1080/17459737.2011.573676

Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates

semantic sentence processing. *Neuropsychologia*, *50*(2), 232–244.

https://doi.org/10.1016/j.neuropsychologia.2011.10.025

Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter

and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693–1708.

https://doi.org/10.1162/jocn.2008.21153

Schröger, E. (1998). Measurement and interpretation of the mismatch negativity. *Behavior Research Methods,*

Instruments and Computers, *30*, 131–145. https://doi.org/10.3758/BF03209423

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical

emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive*

Neuroscience, *18*(8), 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-Brain

responses to rhythmic incongruity. *Cortex*, *45*(1), 80–92. https://doi.org/10.1016/j.cortex.2008.05.014

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Ostergaard, L. (2005). To

musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are

left-lateralized in musicians. *NeuroImage*, *24*(2), 560–564.

https://doi.org/10.1016/j.neuroimage.2004.08.039

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in

music. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7),

2468–2471. https://doi.org/10.1073/pnas.0809035106

Zendel, B. R., & Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive*

Neuroscience, *21*(8), 1488–1498. https://doi.org/10.1162/jocn.2009.21140

Zendel, B. R., Lagrois, M. É., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits

processing of pitch information: The curious case of amusia. *Journal of Neuroscience*, *35*(9),

3815–3824. https://doi.org/10.1523/JNEUROSCI.3766-14.2015

Author Notes

Figure Captions

Figure 1. Samples of the rhythmic sequences: a syntactic-regular rhythmic sequence (Figure 1A) and a syntactic-irregular rhythmic sequence (Figure 1B).

Figure 2. Grand mean ERP waveforms elicited by regular and irregular endings at four electrode sites for (A) musicians and (B) nonmusicians. Gray-shaded areas indicate the time window used for statistical analysis.

Figure 3. Scalp distributions of the irregular-minus-regular difference waves in 100–160 msec and 300–700 msec latency range for (A) musicians and (B) nonmusicians.

Figure 4. Grand mean ERP waveforms elicited by sustained notes in the second, sixth and eighth bars at four electrode sites for the syntactic-irregular sequences in musicians. Gray-shaded areas indicate the time window used for statistical analysis.
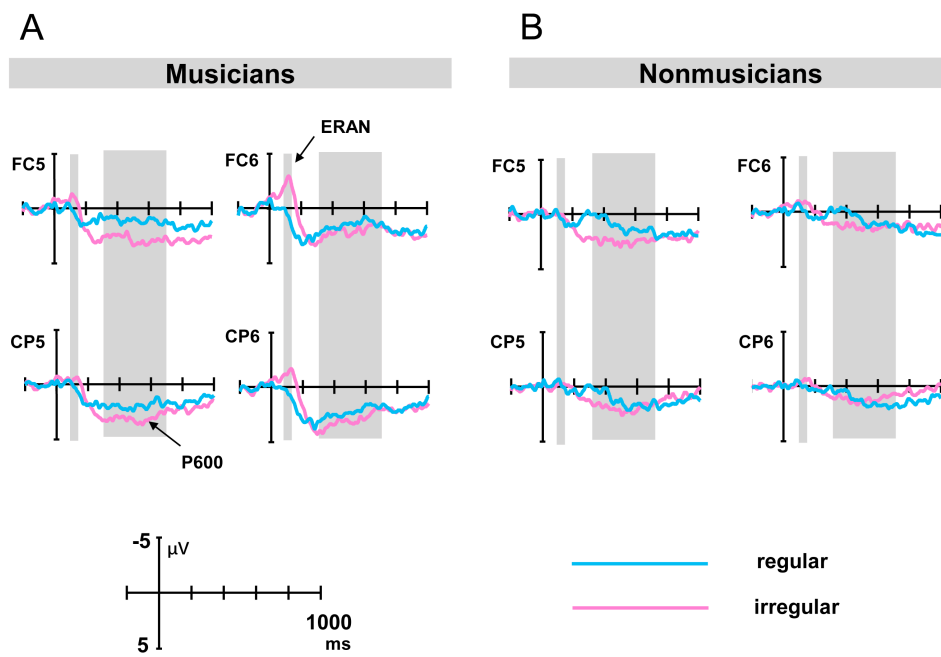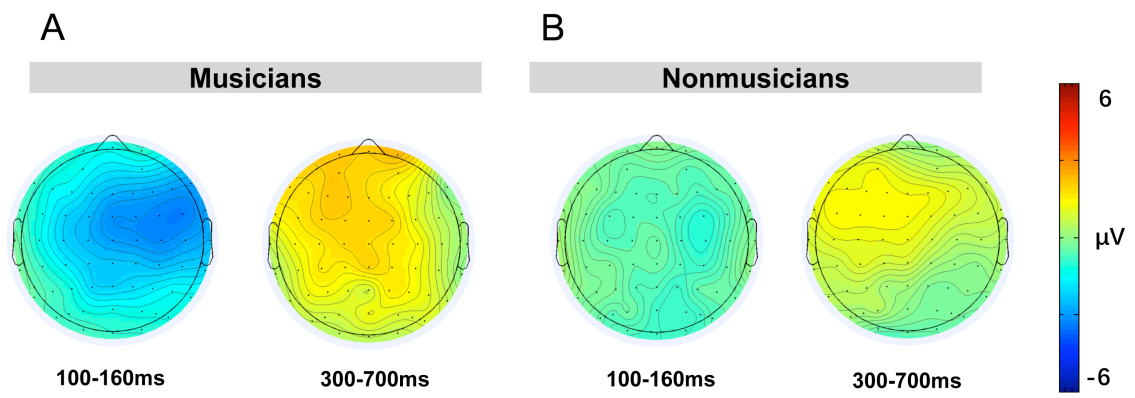
Figure 1

Figure 2

Figure 3

Figure 4

FC5

-3   μV

3

200 ms

FC6

CP5

CP6

sustained tones in the second bar

sustained tones in the sixth bar

sustained tones in the eighth bar