Neuropsychologia 106 (2017) 289–297

Contents lists available at ScienceDirect

Neuropsychologia

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

Neural processing of musical meter in musicians and non-musicians

T. Christina Zhao[a,*], H.T. Gloria Lam[b], Harkirat Sohi[c], Patricia K. Kuhl[a]

[a] Institute for Learning & Brain Sciences (I-LABS), University of Washington, Seattle, WA, United States
[b] Department of Speech and Hearing Sciences, University of Washington, United States
[c] Biomedical & Health Informatics, School of Medicine, University of Washington, United States

ARTICLE INFO

Keywords:
Musical meter
Temporal structure
Tempo
Duple vs. triple meter
Music training
EEG

ABSTRACT

Musical sounds, along with speech, are the most prominent sounds in our daily lives. They are highly dynamic, yet well structured in the temporal domain in a hierarchical manner. The temporal structures enhance the predictability of musical sounds. Western music provides an excellent example: while time intervals between musical notes are highly variable, underlying beats can be realized. The beat-level temporal structure provides a sense of regular pulses. Beats can be further organized into units, giving the percept of alternating strong and weak beats (i.e. metrical structure or meter). Examining neural processing at the meter level offers a unique opportunity to understand how the human brain extracts temporal patterns, predicts future stimuli and optimizes neural resources for processing. The present study addresses two important questions regarding meter processing, using the mismatch negativity (MMN) obtained with electroencephalography (EEG): 1) how tempo (fast vs. slow) and type of metrical structure (duple: two beats per unit vs. triple: three beats per unit) affect the neural processing of metrical structure in non-musically trained individuals, and 2) how early music training modulates the neural processing of metrical structure. Metrical structures were established by patterns of consecutive strong and weak tones (Standard) with occasional violations that disrupted and reset the structure (Deviant). Twenty non-musicians listened passively to these tones while their neural activities were recorded. MMN indexed the neural sensitivity to the meter violations. Results suggested that MMNs were larger for fast tempo and for triple meter conditions. Further, 20 musically trained individuals were tested using the same methods and the results were compared to the non-musicians. While tempo and meter type similarly influenced MMNs in both groups, musicians overall exhibited significantly reduced MMNs, compared to their non-musician counterparts. Further analyses indicated that the reduction was driven by responses to sounds that defined the structure (Standard), not by responses to Deviants. We argue that musicians maintain a more accurate and efficient mental model for metrical structures, which incorporates occasional disruptions using significantly fewer neural resources.

1. Introduction

Music, along with speech, constitute sounds that are universal and ubiquitous in the world's cultures. Both rely on similar acoustic characteristics, such as pitch and timbre variations, to convey information. One important shared characteristic exists in the time domain: both music and speech are highly dynamic, yet temporally structured in a hierarchical manner. The 'Dynamic Attending Theory' and the 'Predictive Coding Theory' posit that the various levels of temporal structure allow the brain to generate predictions and optimize neural resources when processing these temporally dynamic sounds (Jones and Boltz, 1989; Large and Jones, 1999; Snyder and Large, 2005; Vuust and Witek, 2014).

The last decade has seen an increased interest in understanding temporal structure processing. Growing evidence has suggested that the ability to process and track temporal structure may be related to other higher-level cognitive abilities, such as attentional skills (Khalil et al., 2013) and reading (Carr et al., 2014); and that temporal processing skills can change rapidly with early training experiences (Benasich et al., 2014; Tallal et al., 1996; Zhao and Kuhl, 2016).

Western music has a well-characterized hierarchy in the temporal domain, and thus serves as an excellent tool for investigating temporal structure processing at various levels (Fig. 1). While time intervals between notes in rhythmic passages are highly variable, the underlying beats can be realized even when physical notes are absent, providing a beat-level temporal structure (Honing, 2012). The beat-level temporal structure induces a sense of regular or isochronous pulses, which listeners usually feel the urge to synchronize their body movements to,

* Corresponding author. Present address: Institute for Learning & Brain Sciences, University of Washington, BOX 357988, Seattle, WA 98195, United States.
E-mail address: zhaotc@uw.edu (T.C. Zhao).

http://dx.doi.org/10.1016/j.neuropsychologia.2017.10.007
Received 8 June 2017; Received in revised form 1 October 2017; Accepted 3 October 2017
Available online 04 October 2017
0028-3932/

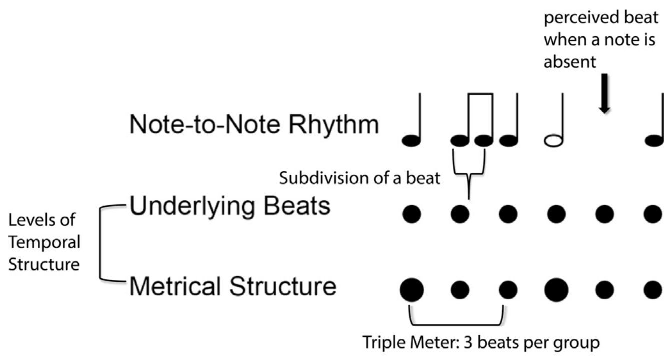

Fig. 1. Illustration of the hierarchy of temporal structure in music.

such as clapping hands and tapping toes (Drake et al., 2000). Further, these isochronous beats can be organized into units of alternating strong and weak beats, establishing a meter-level temporal structure. For example, there are two beats per unit in the duple meter (e.g. marching music: strong-weak-strong-weak), while there are three in the triple meter (e.g. waltz music: strong-weak-weak). This over-arching metrical structure in music provides a sense of grouping for beats, making the strong beats more salient and more predictable (Fitch, 2013). Investigating neural processing of meter-level temporal structure provides a unique window into understanding brain mechanisms associated with the extraction of temporal patterns or groups of beats and the predictions about future beats, in a way that minimizes prediction errors and optimizes neural resources.

The beat-level temporal structure provides the sense of regular pulses that are perceived to occur at equal time intervals. To understand the processing of regular beats, researchers have examined behaviors and neural activities when stimulus streams are presented at either isochronous or random intervals. Compared to randomly timed presentation, stimulus detection thresholds are lower when the stimuli are presented isochronously, and fewer neural resources are involved to process isochronous stimulus streams (Jones et al., 2002; Rohenkohl et al., 2012; Schwartze and Kotz, 2015; Schwartze et al., 2011; van Atteveldt et al., 2015). These results suggest that temporally regular and predictable beats are easier to process than randomly timed stimuli.

Studies have also suggested that the motor system plays a role in tracking the beat-level temporal structure, and that the ability to synchronize body movements to beats requires intricate communication between sensory and motor systems (Arnal, 2012; Patel and Iversen, 2014). Using isochronous stimulus stream vs. randomly timed stimuli, one recent fMRI study observed higher level of activities in the basal ganglia (putamen) for isochronous streams (Geiser et al., 2012). Using stimulus streams varying in their difficulty for beat extraction and perception, other fMRI studies reported that higher levels of activity in the basal ganglia and supplemental motor areas were associated with listening to streams where regular beats were more easily extracted (Grahn and Brett, 2007), and that a stronger perception of beats was related to stronger activity in supplementary motor areas and the premotor cortex across individuals (Grahn and McAuley, 2009). Using EEG, a recent study also demonstrated that participants' beta-band oscillatory activities (highly associated with sensorimotor systems) rebounded after a decrease following a stimulus. The slope of the rebound was associated with stimulus rate only when the stimulus stream was isochronous (not when the stimulus stream was randomly timed), suggesting the endogenous and predictive nature of the rebound (Fujioka et al., 2009, 2012). A recently proposed model suggests that a larger subcortical-cortical network, incorporating the frontal cortex, basal ganglia and cerebellum in addition to the auditory system, is at play for processing beat-level temporal structure (Schwartze and Kotz, 2013).

The meter-level temporal structure involves further organizing beats into groups or units through accenting some beats. While the accents for such perceptual organization are commonly conveyed through acoustic cues in the stimulus, such as intensity (strong-weak), pitch (high-low) or duration (long-short), it has been demonstrated that subjective accents can happen at the perceptual level even when the stimuli are acoustically identical (Brochard et al., 2003; Nozaradan et al., 2011; Potter et al., 2009). For example, using isochronous stimulus streams, individuals responded to deviants more when the deviants occurred at odd positions compared to at even positions in the stream, suggesting that individuals perceptually organized beats into group of two and it was easier to detect deviant at the perceptually stronger locations. The effects were observed to be earlier in the neural responses in musicians than non-musicians (Brochard et al., 2003). Further, when musically trained individuals were instructed to subjectively impose metrical structures (duple or tripe) onto isochronous beats; peaks in their EEG power spectrum were observed at the frequencies corresponding to the imagined meter (Nozaradan et al., 2011).

Other studies examined meter-level temporal structure representation and processes by providing accents in the stimuli. Using short and controlled musical excerpts in different meters, listeners' cortical responses (event-related potentials, or ERPs) to identical sounds were measured and compared when the sounds were at metrically stronger locations in musical passages, or weaker locations (Fitzroy and Sanders, 2015). More specifically, the ERPs to sounds at metrically stronger locations had a more negative N1 peak (a negative peak in event-related potentials at around 100 ms after sound onset) and a more positive P2 peak (a positive peak occurring around 200 ms), compared to ERPs to sounds at metrically weaker locations. This result is interpreted to be in line with the 'Dynamic Attending Theory', hypothesizing that attention is directed to time windows that carry more information (Jones and Boltz, 1989) (i.e. metrically stronger locations). In another study, accents were induced through an increase in intensity in the first part of the isochronous streams. Participants' beta-band modulations to the isochronous tones were measured during the second part of the stream when the acoustic accents disappeared. Beta-band activities were observed to have decreased more after perceptually stronger (or accented) beats than weaker beats even though the sounds were identical (Fujioka et al., 2015). Another commonly used method to study meter-level temporal structure processing is through measuring neural responses to occasional violations to an established meter. Such violation-detection response (e.g. Mismatch Negativity or MMN) can reflect how well the metrical structures were represented. By violating a meter structure at metrically strong and weak positions, non-musicians' MMNs were observed to be earlier when violations occurred at strong metrical positions compared to weak positions, even without attention (Ladinig et al., 2009, 2011). Taken together, research so far has demonstrated that metrically strong beats are processed differently in the brain, in comparison to metrically weak beats.

The current study is designed to examine neural processing of metrical structure from a slightly different perspective: 1) does neural processing of metrically strong beats change with the characteristics of the metrical structures (i.e. Temp: fast vs. slow, Meter Type: duple vs. triple) and 2) whether such neural processing is influenced by early music training experience.

In the current study, we considered the basic unit of metrical structure to be the time interval between strong beats (inter-strong-tone interval). We varied two important parameters of the basic unit to change the metrical structure: tempo and structure type. Tempo was determined by the duration of the inter-strong-tone interval (i.e. shorter time intervals result in a faster tempo). Structure type was determined by the number of beats in each unit (e.g., two beats per unit in duple meter, three beats per unit in triple meter). So far, few studies have specifically examined the effects of these two parameters on meter processing. In one study, participants exhibited shorter latencies in their EEG responses to deviants that violated a duple meter than those that violated a triple meter structure (Abecasis et al., 2005). Using MEG, another study observed differences between processing

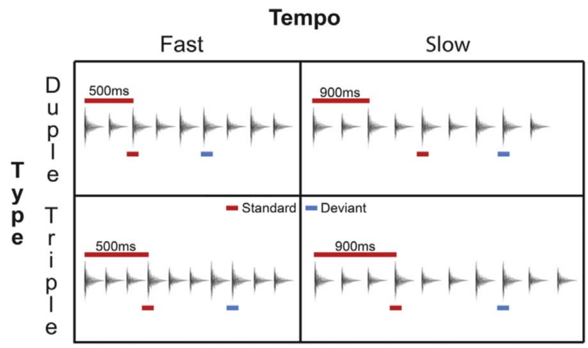

Fig. 2. Illustration of the experimental design. The neural responses to the strong notes were compared across the four conditions. In the fast vs. slow conditions, inter-strong-notes-intervals were 500 ms and 900 ms, respectively. In the duple vs. triple conditions, one or two weak tones were between strong tones, respectively.

downbeats in a duple vs. triple meter to be in cortical as well as subcortical structures (Fujioka et al., 2010).

To address the first research question, we characterized neural responses to a complex tone in participants without musical training (non-musicians). This tone served as the strong tone in four different metrical structures (conditions), with a 2 (Tempo: fast vs. slow) × 2 (Structure Type: duple vs. triple) design (Fig. 2). In each condition, the metrical structure was established by presenting the strong tone along with an attenuated version of it (weak tone) in their corresponding pattern. For example, in the fast/duple condition (Fig. 2. upper left), a strong tone was followed by one weak tone (strong-weak two-tone unit) with an inter-strong-tone interval of 500 ms. Occasionally (15%, every 4–6 units), the metrical structure was violated through the removal of the weak tone(s) between two strong tones. As a result, two strong tones occur consecutively and the second strong tone is considered a Deviant disrupting the metrical structure, by occurring at the point when a weak tone is expected to happen. Using this experimental design, we focused on the tracking and processing of the strong tones, either when they comply with (Standards) or violate (Deviants) the meter.

We first characterized meter processing in these four conditions by examining the event-related potentials (ERPs) to the strong tones, both when they complied with the metrical structure (Standards) and when they violated it (Deviants). However, when the tempo (inter-strong-tone intervals) changed, the note-to-note intervals and number of weak tones inevitably varied across conditions (see Fig. 2). These factors may also affect the ERPs for both Standards and Deviants. For this reason, we then examined the difference in responses to Standards and Deviants (Mismatch Negativity = response to Deviant - response to Standards). Because the Standards and Deviants are acoustically identical, and inter-strong-tone intervals are controlled, the mismatch negativity (MMN) allowed us to specifically target the processes related to the metrical structure violation, subtracting out effects related to response to the acoustics, such as the note-to-note intervals. Although it does not directly measure how meters are represented, the MMN indexes the sensitivity to violations of the predicted strong beats in the target metrical structure, therefore reflecting how well the metrical structures are presented in the brain.

We hypothesized that processing slower tempo would be more challenging due to the increased cognitive load associated with holding longer intervals in working memory and that the inter-strong-tone intervals in fast tempo conditions (500 ms) are also closer to the spontaneous tapping rate in adults (Drake et al., 2000). We also hypothesized that processing triple meter is more challenging because it is less common in Western music, and previous studies suggested that it was easier to detect violations in duple meter than in triple meter (Abecasis et al., 2005; Brochard et al., 2003; Fraisse, 1982).

To address our second question, we examined a group of highly trained musicians to compare to the non-musician group, using the same method. Several previous studies have specifically examined the effects of early music training on meter processing. Using short musical melodies with violations that either disrupted the rhythm but not meter vs. disrupted the meter, it was observed that professional Jazz musicians exhibited a larger and more left-lateralized MMNm as well as a larger P3a, using MEG (Vuust et al., 2009, 2005). These results were replicated in another study that also reported left lateralization in meter processing in musicians (Abecasis et al., 2009). Using tone sequences that follow metrical structures, another set of studies showed better meter violation detection in musicians behaviorally and in MMN (Geiser et al., 2010; Geiser et al., 2009). Therefore, we hypothesized that musicians would exhibit larger MMN in the current study as well, particularly in the more challenging conditions (e.g. slow and triple).

2. Material and methods

2.1. Participants

Two groups of participants were recruited. For the non-musician group, 20 monolingual English-speaking participants were recruited [male = 10, age range: 19–36, M = 22.53 ± 4.69 (SD)]. All participants were non-musicians as defined by a selection criteria used in our previous studies (Zhao and Kuhl, 2015a, 2015b); that is, they have received less than 2 years of private music lessons that ended more than 5 years ago [years of training: 0.37 yrs (M) ± 0.49 (SD)].

For the musician group, 20 additional monolingual English-speaking participants were recruited [male = 10, age range: 20–29, M = 23.61 ± 2.933 (SD)]. All participants were musically trained, that is, they have received more than 8 years of private music lessons that started before the age of 10 [years of training: 12.9 yrs (M) ± 2.53 (SD)] (Zhao and Kuhl, 2015a).

No participant reported any history of hearing, speech or language difficulty. All participants consented and were compensated for their participation.

2.2. Stimuli

The metrical structures with varying tempo and structure type were established by playing strong and weak complex tones in specific sequences in a stream (Fig. 2). The complex tone (Duration: 100 ms, Sampling frequency: 44.1 kHz) was synthesized to have a fundamental frequency of 220 Hz (A3) and a timbre of 'grand piano' timbre along with a 'woodblock' sound in the Overture software (Version 4, Sonic Scores). This complex tone (strong tone) was attenuated by 10 dB in Audacity software (Version 2.0, Sound Forge) to create the weak tone.

2.3. Design and procedures

The experiment was designed to be 2 (Tempo: fast vs. slow) × 2 (Structure Type: Duple vs. Triple) in nature and EEGs were recorded for all 4 conditions in each participant. Order of the condition was randomized across participants. For each condition, 625 trials (or basic units) were presented in which 85% of them followed the metrical structure (Standards) and 15% of them violated the structure (Deviants). Deviant trials were separately by 4–6 Standards.

During an experimental session, participants first completed a short questionnaire regarding their language and music training background. They were then fitted with a 64-channel electrode cap with the conventional 10–20 placement (Electro-Cap International Inc.), with 2 electrodes on the left and right mastoids as references. Three pairs of electro-oculargram (EOG) electrodes (above, below and adjacent to each eye) were used to measure eye movements during the session. The impedances of the electrodes were kept below 15 kΩ. All stimuli were delivered by Stim2 software (Version 4.0), sent from a Dell Optiplex 755 computer to the Stim Audio System, and then to bilateral insert earphones at 80 dB. Both stimulus presentation hardware and software

are part of the Neuroscan system from Compumedics, Inc. Participants listened to the sounds passively while sitting comfortably and watching a silent video of their choice throughout the session in a sound treated booth. All experimental procedures were approved by the University of Washington's Institution Review Board (IRB).

2.4. Data collection and analysis

Continuous EEG was amplified using the SynAmps2™ system and recorded using the Scan (Version 4.5) software at a sampling rate of 1 kHz (Neuroscan, Compumedics). In the preprocessing stage, the EEG data were first offline referenced to the mastoids and down sampled to 250 Hz. Bad channels were visually identified and removed from further analysis. Next, the data were low-pass filtered at 50 Hz and high-pass filtered at 0.1 Hz. Independent components analysis (ICA) was applied to separate and reject artifacts related to eye movements. The data from 6 channels surrounding CZ ('CZ','FCZ','C1','C2','FC1','FC2') were averaged to increase signal-to-noise ratio, as MMN is generally observed to be maximal at CZ electrode All data processing was completed using the EEGLAB software (Delorme and Makeig, 2004) in the MATLAB environment (MathWorks Inc.). Epochs for the strong tones were separately extracted, averaged and baseline corrected for Standards and Deviants in each condition. Trials with voltage values exceeding ± 100 μV were rejected. Data from one non-musician and two musicians were excluded due to excessive noise in the recording.

The Mismatch negativities (MMNs) were subsequently calculated for each condition and for each participant by subtracting responses to standards prior to deviants from responses to deviants, allowing a roughly equal number of standards and deviants in the calculation to minimize the effects of unequal variance.

3. Result

To address our first research question, we present data from the Non-musician group. The averaged responses (ERPs) across participants for the Standards and Deviants of all the conditions are shown in Fig. 3A and B, respectively. Visual inspection of the ERPs for Standards and Deviants (Fig. 3A, B) show clear gradients in the amplitudes in the P2 range (shaded area). The variable note-to-note intervals across conditions was expected to contribute to this gradient, and these results are in line with previous data documenting the effects of inter-stimulus interval (ISI) on N1-P2 responses, with longer ISIs associated with more negative N1 and more positive P2 peaks (Andreou et al., 2015; Crowley and Colrain, 2004; Pereira et al., 2014).

By calculating the difference between responses to Standards and Deviants, the MMN reflects processes that are predominantly associated with metrical violation processing, subtracting out effects related to the acoustic processing and the variable note-to-note intervals (or varied number of weak tones) across conditions. More specifically, the MMN reflects the differences in neural response between a predicted and an unpredicted strong tone, in other words, the error in prediction. We consider the magnitude of the MMN to be indicative of the level of error resulting from the violation of metrical structure during passive listening, that is, the more prominent violations generate bigger errors, resulting in larger MMNs.

The averaged MMN across participants for all conditions is shown in Fig. 4. A negative peak around 150 ms post stimulus onset is present in all conditions and in line with the typical time window of MMN (Naatanen et al., 2007; Winkler et al., 2009). To characterize the effects of Tempo and Type of Temporal Structure on MMN, we first averaged the values in the time window of 125–175 ms for each condition for each participant; we then conducted a repeated measures 2 (Tempo: fast vs. slow) × 2 (Type: duple vs. triple) analysis-of-variance (ANOVA) using the averaged values as the dependent variable. The statistical test suggested a main effect for Tempo, $F(1,18) = 9.40$, $p = 0.007$, partial $\eta^2 = 0.343$, as well as a main effect for Type of Structure, $F(1,18) = 6.83$, $p = 0.018$, partial $\eta^2 = 0.275$; there was no interaction between the two variables, $F(1,18) = 1.55$, $p = 0.229$. Overall, the MMN for faster conditions (mean = − 3.97, 95% CI [− 5.02, − 2.93]) is larger than for slower conditions (mean = − 1.92, 95% CI [− 2.62, − 1.22]) and the MMN for triple conditions (mean = − 3.73, 95% CI [− 4.45, − 3.01]) is larger than for duple conditions (mean = − 2.16, 95% CI [− 3.10, − 1.23]).

To address our second research question, we compare results from Non-musician group to the Musician group. The difference waves for each condition were plotted in Fig. 5A for both groups (left column). Contrary to our prediction, musicians exhibited less negative MMNs in three of the four conditions. The time window between 125 and 175 ms was again averaged and used as the dependent variable for statistical analyses. A mixed 2 (group: musicians vs. non-musicians) × 2 (tempo: fast vs. slow) × 2 (structure type: duple vs. triple) ANOVA revealed a significant main effect for Tempo, $F(1,35) = 9.34$, $p = 0.004$, partial $\eta^2 = 0.211$, a significant main effect for Structure Type, $F(1,35) = 23.44$, $p < 0.001$, partial $\eta^2 = 0.401$, as well as a significant main effect for group, $F(1,35) = 4.82$, $p = 0.035$, partial $\eta^2 = 0.121$. That is, MMNs were more negative in fast tempo (mean = − 3.45, 95% CI [− 4.16, − 2.74]) than slow tempo (mean = − 2.15, 95% CI [− 2.70, − 2.16]), more negative in triple meter conditions (mean = − 3.66, 95% CI [− 4.24, − 3.10]) then duple conditions (mean = − 1.94, 95% CI [− 2.54, − 1.54]). No significant interactions were observed. The results suggested that while Tempo and Type of Structure affected processing of metrical structure similarly in musicians and non-musicians, and there were overall differences between groups.

To further explore this observed difference in MMN between the two groups, the responses to Standards and Deviants were plotted by group in Fig. 5B and C. For the responses to Standards, a consistent difference is clear between the two groups in the P2 range (shaded area). The time window around the peak (averaged between 125 and 175 ms) was used as the dependent variable in another mixed 2 × 2 × 2 ANOVA. The statistical analysis revealed the same pattern: significant main effects for Tempo, $F(1,35) = 58.79$, $p < 0.001$, partial $\eta^2 = 0.627$, a significant main effect for Structure Type, $F(1,35) = 23.69$, $p < 0.001$, partial $\eta^2 = 0.404$, as well as a significant main effect for group, $F(1,35) = 12.97$, $p = 0.001$, partial $\eta^2 = 0.270$. Standards were more positive in slow tempo (mean = − 3.82, 95% CI [3.22, 4.42]) than fast tempo (mean = − 2.01, 95% CI [1.59, 2.41]), more positive in duple meter conditions (mean = 3.6, 95% CI [− 4.24, − 3.10]) than triple conditions (mean = 2.23, 95% CI [1.80, 2.65]). No significant interactions were observed. Musicians exhibited a significantly smaller P2 peak than non-musicians for the strong notes that comply with the temporal structure, regardless of the tempo or type of structure.

The same analysis was then repeated for responses for the Deviants. The analysis revealed a significant interaction between group and Structure type, $F(1,35) = 6.94$, $p = 0.012$, partial $\eta^2 = 0.165$, as well as significant main effects for Tempo, $F(1,35) = 65.27$, $p < 0.001$, partial $\eta^2 = 0.651$, and Structure Type, $F(1,35) = 149.09$, $p < 0.001$, partial $\eta^2 = 0.810$. Again, Deviants were more negative in fast tempo (mean = − 1.44, 95% CI [− 1.99, − 0.89]) than slow tempo (mean = 1.68, 95% CI [1.00, 2.34]), and more negative in triple meter conditions (mean = − 1.44, 95% CI − 1.94, − 0.93]) than duple conditions (mean = − 1.66, 95% CI [1.09, 2.23]). The main effect for group was not significant, $F(1,35) = 1.80$, $p = 0.188$. Further dissection of the interaction shows that for Duple conditions, musicians exhibited a smaller response (mean = 1.02, 95% CI [0.21, 1.83]) than non-musicians (mean = 2.30, 95% CI [1.51, 3.10]), while no difference was found between groups for Triple type.

4. Discussion

4.1. Summaries and interpretations

In the current study, we addressed two research questions that

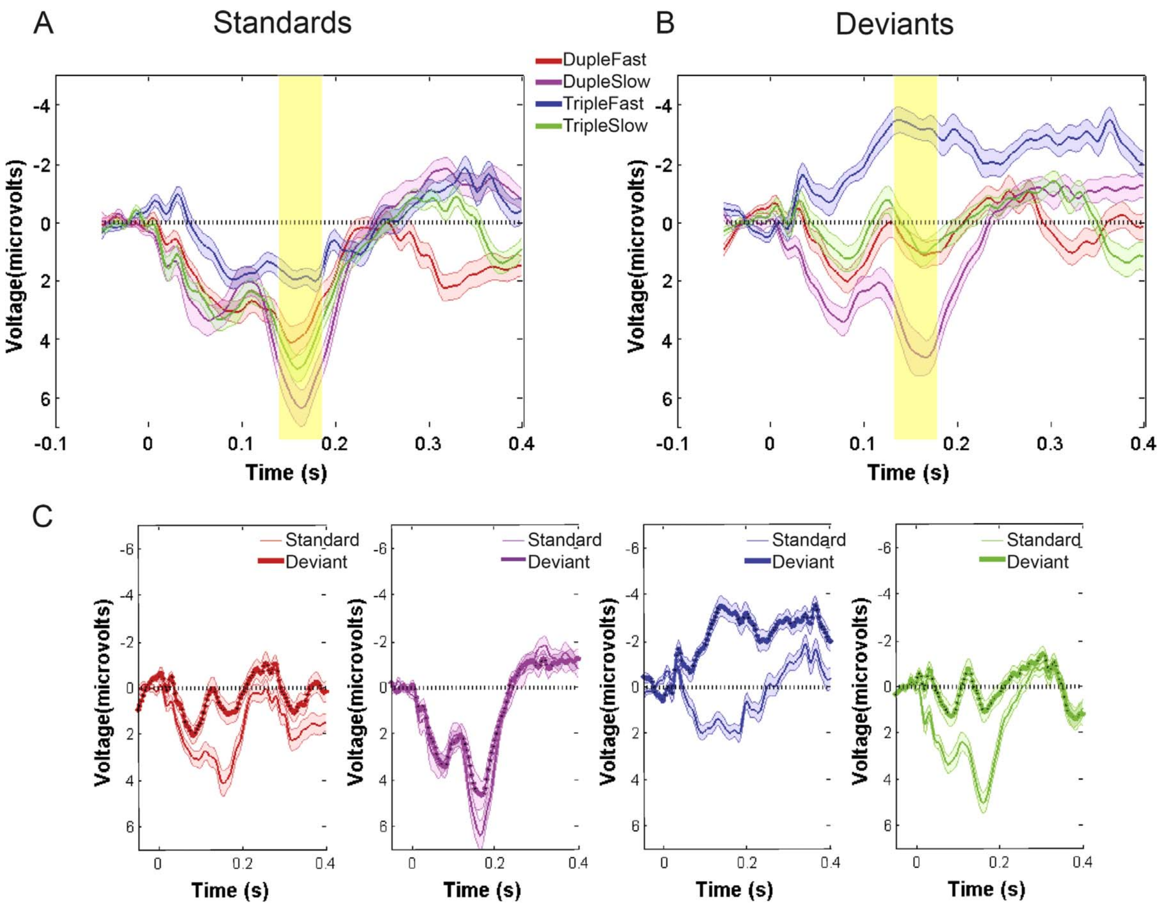

Fig. 3. A) Event-Related Potentials to the standard strong tones averaged across participants with standard error for the 4 conditions. B) Event-Related Potentials to deviant strong tones averaged across participants with standard error for the 4 conditions. C) The same data plotted by conditions. In each subplot (condition), response to Standards (thinner line) is shown with response to Deviants (thicker line).

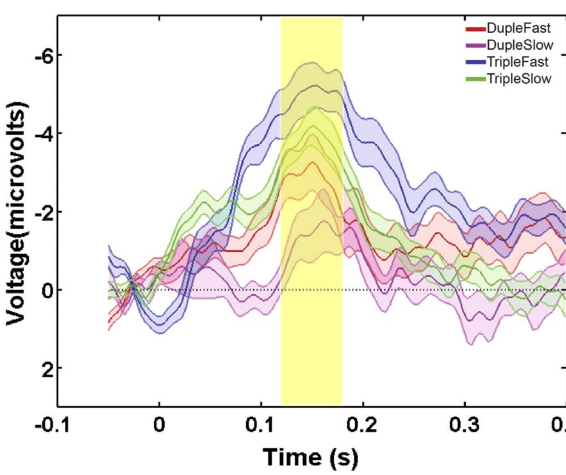

Fig. 4. Difference waves between standards and deviants for the strong tones averaged across participants with standard error for all four conditions.

complemented existing literature on meter-level processing. The first question examines whether neural processing changes with the properties of metrical structures. The second question examines whether neural processing is influenced by early music training experience.

We addressed the first research question by systematically varying the properties of meters (i.e. Tempo and Meter Type) and examining the neural processing in non-musicians. Neural responses to acoustically identical sounds (i.e. strong tones) occurring in compliance (Standards) and in violation (Deviants) to the metrical structures were compared across four different conditions. The difference between responses to Standards and Deviants (MMNs) minimized effects related to acoustic processing (e.g. evoked response to sounds, variable note-to-note intervals and number of weak tones) and reflected the violation detection in the metrical structures, thus providing information on how well the metrical structures are represented in the brain. The results suggested that both Tempo and Type of Temporal structure significantly affected metrical structure violation detection. That is, the MMN is larger in magnitude for faster Tempo conditions and for Triple meter structure conditions. We interpret the results as evidence that violations are more prominent and more noticeable in these conditions, reflected as a larger violation error coded by the neural system (Winkler et al., 2009).

The significant main effect for Tempo supported our hypothesis: the slower conditions with longer units (inter-strong-tone-intervals) resulted in weaker metrical structure representations. This effect could be due to the heavier load on working memory required to maintain the structure as well as the fact that the fast conditions used an interval (500 ms) that is closer to adults spontaneous tapping rate, or internal reference rate (Drake et al., 2000). Therefore, the violation errors to the metrical structure are not as prominent in the slow conditions. In contrast, the main effect for Structure Type did not support our hypothesis: that is, we found larger MMNs for the triple meter conditions than the duple meter conditions. We discuss two factors that possibly contributed to this finding. First, in comparison to previous studies examining the difference between duple and triple meters, our current study controlled for the duration and number of inter-strong-tone intervals instead of the note-to-note intervals. In previous studies, by controlling for note-to-note intervals, the triple conditions thus have longer units (inter-strong-tone intervals) and fewer metrically strong

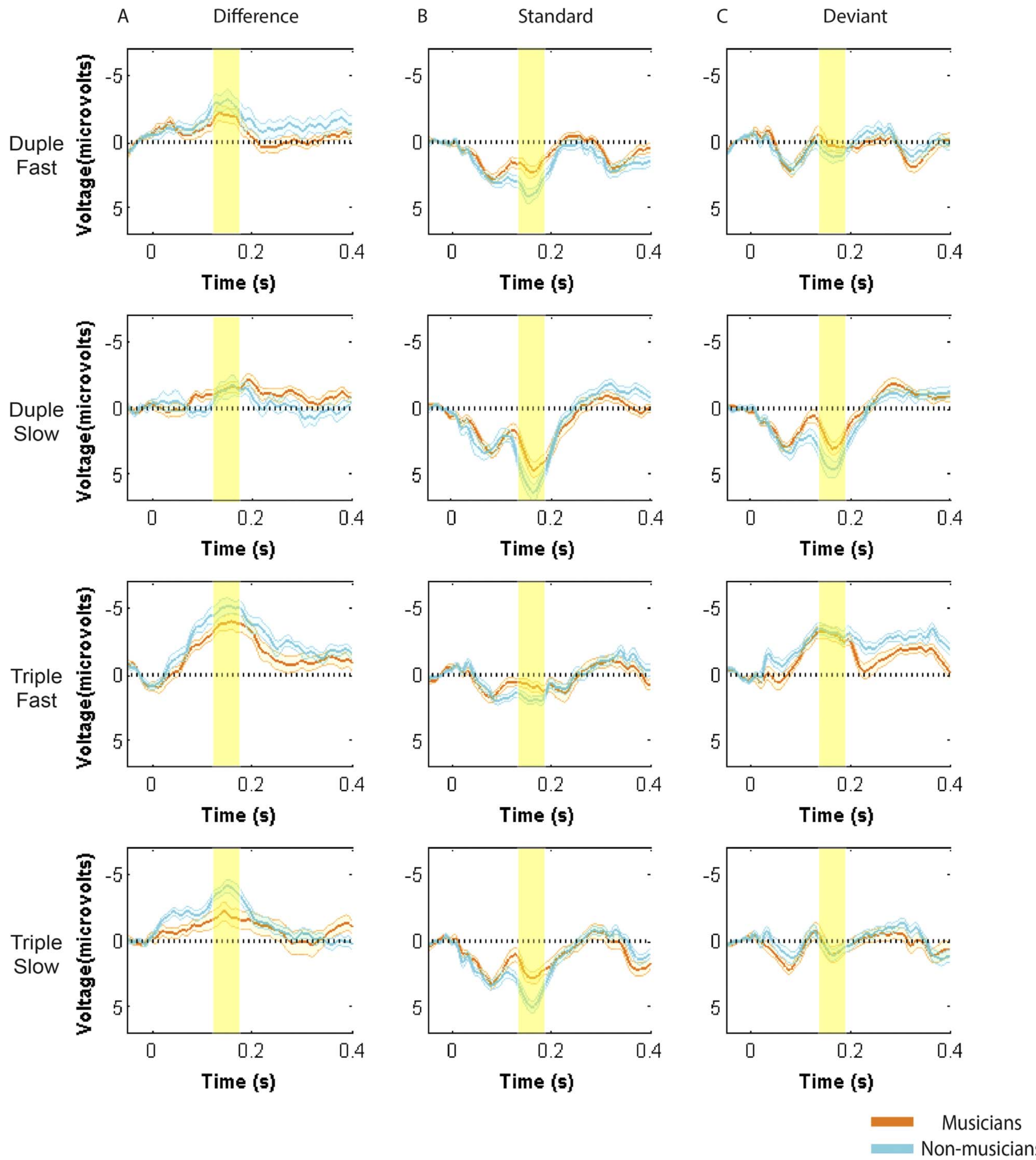

Fig. 5. A) Difference waves for all four conditions between musicians and non-musicians B) Responses to Standards for all four conditions between musicians and non-musicians. C) Responses to Deviants for all four conditions for musicians and non-musicians.

beats. That may contribute to previously observed difference that duple meters are easier to process (Abecasis et al., 2005; Fujioka et al., 2010). We provide further discussion on controlling for inter-strong-tone intervals vs. note-to-note intervals in the 'Limitation' section. The second factor that may have contributed to the current results concerns the location of violation. In the Standard trials of triple meter structure (i.e. group of 3), we observed that the neural response to the last tone (second yellow shaded region, Fig. 6) in the 3-tone unit is significantly larger than the response to the middle tones (first yellow shaded region, Fig. 6), even though the two tones are identical in terms of their acoustic information (Fig. 6 plots the Triple Slow condition as an example). This phenomenon has been documented in previous research, which showed differences between the two weak tones in a triple meter structure as well as the later two syllables in a three-syllable unit considered to be a word (Bosseler et al., 2016; Schaefer et al., 2011). We speculate that the perceptual weight of the second tone is therefore the weakest. By removing all weak tones between two strong tones, the tone prior to the Deviant is ensured to be identical (i.e. a strong tone) across conditions. At the same time, in the triple meter conditions, the Deviant violated the metrical structure at the perceptually weakest tone position, resulting in a more prominent violation. Further research is warranted to compare the perceptual strength of all

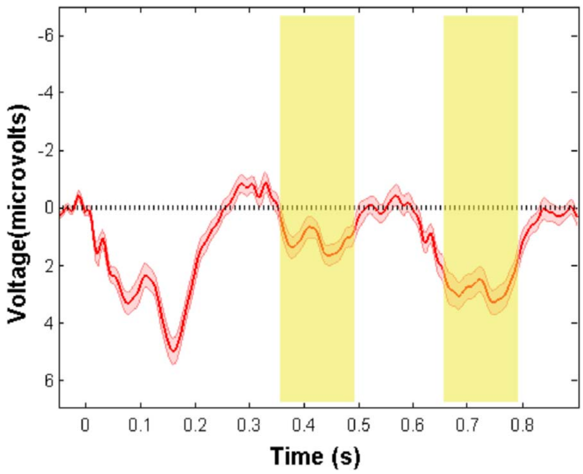

Fig. 6. Event-related potentials to the 3 tones (Strong-weak-weak) in a unit (Standards) in the Triple slow condition. The response to the last tone (second yellow shaded area) is significantly stronger than the middle tone (first yellow shaded area) even though their acoustics are identical. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

weak tones across meters.

To address the second question, we tested highly trained musicians using the same methods and compared their data to the non-musicians. Contrary to our predictions, musicians showed smaller MMNs on all of the conditions except for Duple Slow, driving a significant main effect of group. We further explored this group difference in MMN by directly comparing the two groups' responses to Standards and Deviants. Our analyses revealed a significantly smaller P2 response for the Standards for all conditions in musicians, while no difference was observed for the Deviants between groups except for in the Duple Slow condition.

The P2 response is an ERP component that has been widely documented, but its function remains unclear. Researchers have suggested that the level of cognitive load during processing affects P2 responses (Crowley and Colrain, 2004). Previous studies reported that when presented with tones of different spectral complexity, P2 is larger for more complex sounds and musicians on average demonstrated a larger P2 response (Shahin et al., 2003, 2005). On the other hand, it has also been reported that the P2 amplitude decreased for musicians over time in a single experimental session for tones of various pitch, reflecting a rapid learning effect. This effect was not observed in non-musicians (Seppänen et al., 2012).

Based on previous data, we interpreted our results as follows: using metrical structures that are highly common in Western culture and presenting them in an extremely simple manner (no pitch or timbre variation and continuous stream) may have allowed the musicians to extract and maintain a very accurate representation of the metrical structure (i.e. a representation that minimized prediction errors by taking the occasional violations into account), using reduced cognitive resources. On the other hand, within each group, the more difficult temporal structure is related to larger P2 responses, requiring more cognitive resources for processing. When occasional violations (every 4–6 units) are represented and accounted for in musicians' mental models of the metrical structure, prediction errors (reflected by MMN) become minimized and are therefore less prominent.

We speculate that the level of complexity in the stimulus stream carrying the metrical structures influenced the accuracy and the maintenance of the mental model of the metrical structure. In two studies, the metrical structure was delivered in a complex manner with pitch and timbre variations (Vuust et al., 2009, 2005). At the same time, these studies used short tone sequences/short musical melodies, in comparison to our continuous presentation. Participants had to re-extract metrical structure in each trial in previous studies, making it a harder task to extract and maintain the representation of the metrical structure. Thus, the representation itself may be less accurate. In contrast, the stimuli in the current study are continuous with mono-tone and mono-timbre with the same type of occasional violations, thus it becomes extremely easy for highly trained musicians to maintain a more accurate representation that tolerates the occasional violations, in an efficient manner. Future research is warranted to systematically vary the complexity of the stimuli that are carrying the metrical structure and examine the accuracy and maintenance of the metrical structure in listeners, given various complexities.

4.2. Limitations

Our current study is limited by several factors. First, in establishing our meters, we explicitly controlled the inter-strong-tone intervals to be identical, therefore introducing variable note-to-note intervals and number of weak tones. This was different from previous studies (Abecasis et al., 2005; Fujioka et al., 2010) that used designs that kept the note-to-note intervals (and number of tones) equal across conditions. Our specific design was chosen because our research questions only focused on metrical structure violation at the strong beat locations. Keeping the inter-strong-tone intervals constant minimized other effects, for example, if note-to-note intervals and total number of tones are constant, the intervals between two strong tones would be longer and there would be fewer strong tones in triple meter than duple. Any difference observed between the two meters may be confounded by the difference in their inter-strong-tone intervals. The inter-strong-tone interval and note-to-note interval issue poses difficulty in designs that are hard to resolve in a single study. Future studies are warranted to further address the variable note-to-note issue in the current study; for example, additional conditions where note-to-note intervals are controlled can provide additional control for explaining our current results.

Future studies are warranted to also control for potential effects related to counting or sequence learning. A recent study elegantly isolated the effect of sequence learning from meter processing by presenting the same sequence of alternating tones either bounded by a metrical structure or temporally jittered. The results suggested that violation detection at the strong beat position were better when bounded by a metrical structure than when in a temporally-jittered sequence, demonstrating effects specific to meter processing (Bouwer et al., 2016).

Lastly, the current study may be limited by the level of control we have on participants' overall attention level. Though all participants were watching a silent video of their choice with the instruction to ignore the sounds from the headphones, musically trained individuals may be allocating a different amount of attentional resources unconsciously to the sounds than the non-musicians. Future research will explore paradigms that exert higher levels of control on attention to further examine the effects of attention on temporal structure processing.

4.3. General discussion

The prominent sounds we hear on a daily basis (e.g. music and speech) are highly dynamic but also well-structured in the time domain. These temporal structures, or patterns, constitute an important aspect of auditory processing, allowing the brain to predict future stimuli and optimize neural resources for processing. The current study demonstrated that the neural processing of metrical structure is affected not only by the parameters that characterize the structure (Tempo and Structure Type), but is also modulated by the early music training experiences of the listeners. In highly trained musicians, the responses are significantly reduced in comparison to non-musicians, and the reductions are more prominent in less challenging conditions. We speculate that musicians are able to maintain a more accurate mental representation using less cognitive resources, reflecting higher levels of processing efficiency.

Processing complex auditory stimuli with increasing efficiency reflects an important aspect of learning, which has been reported in areas such as phonetic learning (Zhang et al., 2009). In the early phases of development, an explosion in the number of neurons first establishes many connections to accommodate various types of information processing. A long period of pruning then follows to make relevant pathways more efficient while reducing pathways that become irrelevant (Ponton et al., 2000; Wunderlich and Cone-Wesson, 2006). In the realm of metrical processing, our previous study with 9-month-old infants using similar stimuli demonstrated that one-month of musical intervention in the laboratory at a very early stage of development enhanced infants' ability to process triple meter structure, a difficult meter for infants at that age (Zhao and Kuhl, 2016). In the present study, we show that adults who had experienced extensive music training during early development may be processing temporal structure in a more efficient and accurate manner in adulthood. Cross-sectional or longitudinal studies that sample at multiple time points during development are warranted to help elucidate a full picture of experience-related effects with different durations, intensities of musical experience at different developmental stages.

Acknowledgement

The research was funded by private gifts to the I-LABS Ready Mind Project and funds from Mr. Bard Richmond. We thank Dr. Alexander Khalil and Dr. Steven Morrison for comments on an earlier draft.

References

Abecasis, D., Brochard, R., Del Río, D., Dufour, A., Ortiz, T., 2009. Brain lateralization of metrical accenting in musicians. Ann. N.Y. Acad. Sci. 1169 (1), 74–78.

Abecasis, D., Brochard, R., Granot, R., Drake, C., 2005. Differential brain response to metrical accents in isochronous auditory sequences. Music Percept.: Interdiscip. J. 22 (3), 549–562.

Andreou, L.V., Griffiths, T.D., Chait, M., 2015. Sensitivity to the temporal structure of rapid sound sequences - an MEG study. Neuroimage 110, 194–204. http://dx.doi.org/10.1016/j.neuroimage.2015.01.052.

Arnal, L.H., 2012. Predicting "when" using the motor system's beta-band oscillations. Front. Hum. Neurosci. 6. http://dx.doi.org/10.3389/fnhum.2012.00225.

Benasich, A.A., Choudhury, N.A., Realpe-Bonilla, T., Roesler, C.P., 2014. Plasticity in developing brain: active auditory exposure impacts prelinguistic acoustic mapping. J. Neurosci. 34 (40), 13349–13363. http://dx.doi.org/10.1523/jneurosci.0972-14.2014.

Bosseler, A.N., Teinonen, T., Tervaniemi, M., Huotilainen, M., 2016. Infant directed speech enhances statistical learning in newborn infants: an ERP study. PLoS One 11 (9), e0162177.

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. http://dx.doi.org/10.1016/j.neuropsychologia.2016.02.018.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., Drake, C., 2003. The "Ticktock" of our internal clock. Psychol. Sci. 14 (4), 362–366. http://dx.doi.org/10.1111/1467-9280.24441.

Carr, K.W., White-Schwoch, T., Tierney, A.T., Strait, D.L., Kraus, N., 2014. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. Proc. Natl. Acad. Sci. USA 111 (40), 14559–14564. http://dx.doi.org/10.1073/pnas.1406219111.

Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115 (4), 732–744. http://dx.doi.org/10.1016/j.clinph.2003.11.021.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134, 9–21.

Drake, C., Jones, M.R., Baruch, C., 2000. The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. Cognition 77 (3), 251–288. http://dx.doi.org/10.1016/s0010-0277(00)00106-2.

Fitch, W.T., 2013. Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. Front. Syst. Neurosci. 7. http://dx.doi.org/10.3389/fnsys.2013.00068.

Fitzroy, A.B., Sanders, L.D., 2015. Musical meter modulates the allocation of attention across time. J. Cogn. Neurosci.

Fraisse, P., 1982. Rhythm and tempo. In: Deutsch, D. (Ed.), The Psychology of Music. Academy Press, New York.

Fujioka, T., Ross, B., Trainor, L.J., 2015. Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. J. Neurosci. 35 (45), 15187–15198. http://dx.doi.org/10.1523/jneurosci.2397-15.2015.

Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2009. Beta and gamma rhythms in human auditory cortex during musical beat processing. In: DallaBella, S., Kraus, N., Overy, K., Pantev, C., Snyder, J.S., Tervaniemi, M., Tillmann, B., Schlaug, G. (Eds.), Neurosciences and Music III: Disorders and Plasticity Vol. 1169. Blackwell Publishing, Oxford, pp. 89–92.

Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2012. Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J. Neurosci. 32 (5), 1791–1802. http://dx.doi.org/10.1523/jneurosci.4107-11.2012.

Fujioka, T., Zendel, B.R., Ross, B., 2010. Endogenous neuromagnetic activity for mental hierarchy of timing. J. Neurosci. 30 (9), 3458–3466. http://dx.doi.org/10.1523/jneurosci.3086-09.2010.

Geiser, E., Notter, M., Gabrieli, J.D.E., 2012. A corticostriatal neural system enhances auditory perception through temporal context processing. J. Neurosci. 32 (18), 6177–6182. http://dx.doi.org/10.1523/jneurosci.5153-11.2012.

Geiser, E., Sandmann, P., Jancke, L., Meyer, M., 2010. Refinement of metre perception - training increases hierarchical metre processing. Eur. J. Neurosci. 32 (11), 1979–1985. http://dx.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. http://dx.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. J. Cogn. Neurosci. 19 (5), 893–906. http://dx.doi.org/10.1162/jocn.2007.19.5.893.

Grahn, J.A., McAuley, J.D., 2009. Neural bases of individual differences in beat perception. Neuroimage 47 (4), 1894–1903. http://dx.doi.org/10.1016/j.neuroimage.2009.04.039.

Honing, H., 2012. Without it no music: beat induction as a fundamental musical trait. In: Neurosciences and Music Iv: Learning and Memory, vol. 1252, Oxford: Blackwell Science Publ., pp. 85–91.

Jones, M.R., Boltz, M., 1989. Dynamic attending and responses to time. Psychol. Rev. 96 (3), 459–491. http://dx.doi.org/10.1037//0033-295x.96.3.459.

Jones, M.R., Moynihan, H., MacKenzie, N., Puente, J., 2002. Temporal aspects of stimulus-driven attending in dynamic arrays. Psychol. Sci. 13 (4), 313–319. http://dx.doi.org/10.1111/1467-9280.00458.

Khalil, A.K., Minces, V., McLoughlin, G., Chiba, A., 2013. Group rhythmic synchrony and attention in children. Front. Psychol. 4, 564. http://dx.doi.org/10.3389/fpsyg.2013.00564.

Ladinig, O., Honing, H., Háàden, G., Winkler, I., 2009. Probing attentive and preattentive emergent meter in adult listeners without extensive music training. Music Percept.: Interdiscip. J. 26 (4), 377–386. http://dx.doi.org/10.1525/mp.2009.26.4.377.

Ladinig, O., Honing, H., Haden, G.P., Winkler, I., 2011. Erratum. Music Percept.: Interdiscip. J. 28 (4). http://dx.doi.org/10.1525/mp.2011.28.4.444. (444-444).

Large, E.W., Jones, M.R., 1999. The dynamics of attending: how people track time-varying events. Psychol. Rev. 106 (1), 119–159. http://dx.doi.org/10.1037/0033-295x.106.1.119.

Naatanen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin. Neurophysiol. 118 (12), 2544–2590. http://dx.doi.org/10.1016/j.clinph.2007.04.026.

Nozaradan, S., Peretz, I., Missal, M., Mouraux, A., 2011. Tagging the neuronal entrainment to beat and meter. J. Neurosci. 31 (28), 10234–10240. http://dx.doi.org/10.1523/jneurosci.0411-11.2011.

Patel, A.D., Iversen, J.R., 2014. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Front. Syst. Neurosci. 8, 57. http://dx.doi.org/10.3389/fnsys.2014.00057.

Pereira, D.R., Cardoso, S., Ferreira-Santos, F., Fernandes, C., Cunha-Reis, C., Paiva, T.O., Marques-Teixeira, J., 2014. Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. Int. J. Psychophysiol. 94 (3), 311–318. http://dx.doi.org/10.1016/j.ijpsycho.2014.09.012.

Ponton, C.W., Eggermont, J.J., Kwong, B., Don, M., 2000. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. Clin. Neurophysiol. 111 (2), 220–236. http://dx.doi.org/10.1016/s1388-2457(99)00236-9.

Potter, D.D., Fenwick, M., Abecasis, D., Brochard, R., 2009. Perceiving rhythm where none exists: event-related potential (ERP) correlates of subjective accenting. Cortex 45 (1), 103–109. http://dx.doi.org/10.1016/j.cortex.2008.01.004.

Rohenkohl, G., Cravo, A.M., Wyart, V., Nobre, A.C., 2012. Temporal expectation improves the quality of sensory information. J. Neurosci. 32 (24), 8424–8428. http://dx.doi.org/10.1523/jneurosci.0804-12.2012.

Schaefer, R.S., Vlek, R.J., Desain, P., 2011. Decomposing rhythm processing: electroencephalography of perceived and self-imposed rhythmic patterns. Psychol. Res. 75 (2), 95–106. http://dx.doi.org/10.1007/s00426-010-0293-4.

Schwartze, M., Kotz, S.A., 2013. A dual-pathway neural architecture for specific temporal prediction. Neurosci. Biobehav. Rev. 37 (10), 2587–2596. http://dx.doi.org/10.1016/j.neubiorev.2013.08.005.

Schwartze, M., Kotz, S.A., 2015. The timing of regular sequences: production, perception, and covariation. J. Cogn. Neurosci. 27 (9), 1697–1707. http://dx.doi.org/10.1162/jocn_a_00805.

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., Kotz, S.A., 2011. Temporal regularity effects on pre-attentive and attentive processing of deviance. Biol. Psychol. 87 (1), 146–151. http://dx.doi.org/10.1016/j.biopsycho.2011.02.021.

Seppänen, M., Hämäläinen, J., Pesonen, A.-K., Tervaniemi, M., 2012. Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. Front. Hum. Neurosci. 6 (43). http://dx.doi.org/10.3389/fnhum.2012.00043.

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E., 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J. Neurosci. 23 (13), 5545–5552.

Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B., 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport

16 (16), 1781–1785. http://dx.doi.org/10.1097/01.wnr.0000185017.29316.63.
Snyder, J.S., Large, E.W., 2005. Gamma-band activity reflects the metric structure of rhythmic tone sequences. Cogn. Brain Res. 24 (1), 117–126. http://dx.doi.org/10.1016/j.cogbrainres.2004.12.014.
Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X.Q., Nagarajan, S.S., Merzenich, M.M., 1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science 271 (5245), 81–84. http://dx.doi.org/10.1126/science.271.5245.81.
van Atteveldt, N., Musacchia, G., Zion-Golumbic, E., Sehatpour, P., Javitt, D.C., Schroeder, C., 2015. Complementary fMRI and EEG evidence for more efficient neural processing of rhythmic vs. unpredictably timed sounds. Front. Psychol. 6, 1663. http://dx.doi.org/10.3389/fpsyg.2015.01663.
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. http://dx.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. http://dx.doi.org/10.1016/j.neuroimage.2004.08.039.
Vuust, P., Witek, M.A.G., 2014. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. Front. Psychol. 5, 14. http://dx.doi.org/10.3389/fpsyg.2014.01111.
Winkler, I., Denham, S.L., Nelken, I., 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn. Sci. 13 (12), 532–540. http://dx.doi.org/10.1016/j.tics.2009.09.003.
Wunderlich, J.L., Cone-Wesson, B.K., 2006. Maturation of CAEP in infants and children: a review. Hear. Res. 212 (1–2), 212–223. http://dx.doi.org/10.1016/j.heares.2005.11.008.
Zhang, Y., Kuhl, P.K., Imada, T., Iverson, P., Pruitt, J., Stevens, E.B., Nemoto, I., 2009. Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. Neuroimage 46 (1), 226–240. http://dx.doi.org/10.1016/j.neuroimage.2009.01.028.
Zhao, T.C., Kuhl, P.K., 2015a. Effect of musical experience on learning lexical tone categories. J. Acoust. Soc. Am. 137 (3), 1452–1463. http://dx.doi.org/10.1121/1.4913457.
Zhao, T.C., Kuhl, P.K., 2015b. Higher-level linguistic categories dominate lower-level acoustics in lexical tone processing. J. Acoust. Soc. Am. 138 (2), EL133–EL137. http://dx.doi.org/10.1121/1.4927632.
Zhao, T.C., Kuhl, P.K., 2016. Musical intervention enhances infants' neural processing of temporal structure in music and speech. Proc. Natl. Acad. Sci. USA 113 (19), 5212–5217. http://dx.doi.org/10.1073/pnas.1603984113.