

# Rhythm Complexity Modulates Behavioral and Neural Dynamics During Auditory–Motor Synchronization

Brian Mathias<sup>1,2\*</sup>, Anna Zamm<sup>1,3\*</sup>, Pierre G. Gianferrara<sup>1,4</sup>,  
Bernhard Ross<sup>5</sup>, and Caroline Palmer<sup>1</sup>

## Abstract

■ We addressed how rhythm complexity influences auditory–motor synchronization in musically trained individuals who perceived and produced complex rhythms while EEG was recorded. Participants first listened to two-part auditory sequences (Listen condition). Each part featured a single pitch presented at a fixed rate; the integer ratio formed between the two rates varied in rhythmic complexity from low (1:1) to moderate (1:2) to high (3:2). One of the two parts occurred at a constant rate across conditions. Then, participants heard the same rhythms as they synchronized their tapping at a fixed rate (Synchronize condition). Finally, they tapped at the same fixed rate (Motor condition). Auditory feedback from their taps was present in all conditions. Behavioral effects of rhythmic complexity were evidenced in all tasks; detection of

missing beats (Listen) worsened in the most complex (3:2) rhythm condition, and tap durations (Synchronize) were most variable and least synchronous with stimulus onsets in the 3:2 condition. EEG power spectral density was lowest at the fixed rate during the 3:2 rhythm and greatest during the 1:1 rhythm (Listen and Synchronize). ERP amplitudes corresponding to an N1 time window were smallest for the 3:2 rhythm and greatest for the 1:1 rhythm (Listen). Finally, synchronization accuracy (Synchronize) decreased as amplitudes in the N1 time window became more positive during the high rhythmic complexity condition (3:2). Thus, measures of neural entrainment corresponded to synchronization accuracy, and rhythmic complexity modulated the behavioral and neural measures similarly. ■

## INTRODUCTION

Many behaviors require the temporal coordination of one's actions with perceived auditory information, from dance (Brown & Parsons, 2008), to athletics (Bood, Nijssen, Van Der Kamp, & Roerdink, 2013), to music-making (Wing, Endo, Bradbury, & Vorberg, 2014). When dancing to music, for example, one must coordinate the timing of body movements with a perceived auditory rhythm, a temporally regular acoustic pattern (Miura, Kudo, & Nakazawa, 2013; Thaut, 2013). The temporal alignment of one's movement with the frequency and phase of auditory rhythms can be described as “auditory–motor synchronization.” Musical auditory–motor synchronization requires the simultaneous perception of auditory rhythms and efficient coordination of movement with those rhythms (for a review, see Repp & Su, 2013). A major question in cognitive neuroscience is how interactions between auditory and motor processes give rise to accurate auditory–motor synchronization.

Emerging evidence suggests that perception of auditory rhythms is accompanied by neural forms of entrainment, defined here as the process by which oscillations couple with, or alter their period in response to, another (intrinsic

oscillation or a stimulus rhythm (Haken, Kelso, & Bunz, 1985). Neural entrainment can arise when rhythmic fluctuations of electrical brain activity or neural oscillations arise from synchronous excitability in networks of functionally connected neurons (Buzsáki & Draguhn, 2004), which become coupled with (adapt their period in response to) acoustic signals. Neural entrainment with frequency components of auditory rhythms is evidenced in EEG measures of neural responses that occur at positions of occasional omitted tones in rhythmic auditory sequences (Snyder & Large, 2005) and enhanced amplitudes of neural oscillations at the periodicity of a perceived beat in auditory tone sequences that did not contain energy at that periodicity (Fujioka, Ross, & Trainor, 2015; Nozaradan, Peretz, & Mouraux, 2012; Nozaradan, Peretz, Missal, & Mouraux, 2011).

The coupling of neural oscillations with acoustic rhythms can enhance perceptual processing of stimulus events (Large & Palmer, 2002; Engel, Fries, & Singer, 2001; Large & Jones, 1999). EEG studies have shown that mid-latency ERPs are enhanced in response to acoustic events that receive enhanced perceptual processing, such as those to which individuals voluntarily attend (Sowman, Kuusik, & Johnson, 2012; Hillyard, Hink, Schwent, & Picton, 1973). These effects have been observed in the N1 component, a negative-going ERP component that peaks around 100 msec after tone onsets (Nobre & van Ede, 2018; Lange, Rösler, & Röder, 2003; Näätänen & Winkler, 1999). N1 amplitudes are typically measured as the mean amplitude across a

<sup>1</sup>McGill University, <sup>2</sup>Max Planck Institute for Human Cognitive and Brain Science, <sup>3</sup>Central European University, Budapest, Hungary, <sup>4</sup>Carnegie Mellon University, <sup>5</sup>Rotman Research Institute, Toronto, ON, Canada

\*Joint first authors.

post-stimulus time window or as the negative peak within a post-stimulus time window. Some alternatives to these window-based measures involve defining N1 amplitude as the voltage difference between the most prominent peak in the N1 time window and the peak of a neighboring ERP such as the P1 (i.e., peak-to-peak measurement). Although these alternative peak-to-peak metrics may in some cases be able to disambiguate multiple overlapping ERPs, the window-based approach is less susceptible to artifactual peaks in the data—which can arise from filtering artifacts and other noise sources (Woodman, 2010)—and is more widely used in the literature on auditory rhythm perception. Moreover, P1, N1, and P2 components are often analyzed separately because they are thought to reflect different functions.

More negative amplitudes within post-stimulus time windows associated with the N1 have been observed in response to tones aligned with metrically strong beats compared to metrically weak beats during the perception of short melodies (Fitzroy & Sanders, 2015) and to sounds that are accented by increased intensity in short rhythmic sequences (Schaefer, Vlek, & Desain, 2011). If individuals allocate attentional resources to tones that occur at rhythmically salient frequencies in auditory sequences, then amplitudes within an N1-related time window elicited by these attended tones should be larger than for tones which occur at less-attended frequencies. A final effect on amplitudes within the timeframe of the N1 component is repetition: Repeated sounds tend to elicit a smaller N1 response than novel, non-repeated sounds. This phenomenon may arise from refraction in auditory circuits (Budd, Barry, Gordon, Rennie, & Michie, 1998), which is the temporal interval across which a given neural system returns to baseline excitability, or from sensory memory updating (for a review, see Näätänen & Picton, 1987).

N1 amplitudes are reduced (more positive) in response to self-generated relative to externally generated sounds, possibly because of motor-induced suppression of auditory processing (Horváth, Maess, Baess, & Tóth, 2012). Cortical motor regions may generate templates of sounds that we intend to produce (Bays & Wolpert, 2007); these templates are accessible to sensory memory and are subtracted from actual sensory input during production, reflected in suppressed amplitudes within a time window surrounding the N1 (SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger, 2013). Therefore, N1 responses may distinguish between sounds that one produces and sounds that one perceives during auditory–motor synchronization. We test here whether larger amplitudes within the typical timeframe of the N1 are elicited in response to the frequencies that participants synchronize with (hear) but do not produce, than to the frequencies that participants produce.

Recent research suggests that production modulates not only auditory–motor ERPs but also oscillatory brain responses. Larger neural oscillatory responses have been observed when individuals tap along with a rhythm compared to only listening to the rhythm (Nozaradan, Schönwiesner, Caron-Desrochers, & Lehmann, 2016). Moving one’s body

at a specific frequency to a rhythmic auditory sequence can enhance amplitudes of neural oscillations at that frequency while subsequently listening to the same auditory sequence (Chemin, Mouraux, & Nozaradan, 2014). Neural oscillations during rhythm perception may also predict synchronization accuracy during rhythmic production: Stronger oscillations at the frequency of a perceived rhythm are associated with greater temporal prediction during sensorimotor synchronization (Nozaradan, Peretz, & Keller, 2016).

An open question is how neural oscillations moderate perceptual processing of complex ecological auditory rhythms such as those occurring in multipart music. One approach comes from the dynamical systems framework of rhythmic entrainment (Large, Herrera, & Velasco, 2015; Strogatz, 2001), which describes mathematically how coupling arises between oscillators with different frequencies. The stability of two rhythms (or periodic oscillations) is a function of the ratio between their frequencies; rhythms that form a simple integer ratio relationship (such as 1:1), referred to here as “simple rhythms,” achieve more stability than rhythms that form a complex integer ratio (such as 3:2), referred to here as “complex rhythms” (Glass & Mackey, 1988). The relationships between the rhythm ratios can be described by a Farey tree, which defines the regions of stability for two rhythms based on their frequency ratio (Kelso, 1991; Schroeder, 1991). Dynamical models offer predictions for how neural oscillations respond to simple versus complex auditory rhythms, namely, the stability of neural oscillations responding to auditory rhythms based on their rhythmic ratios. In turn, more stable neural oscillations may enhance perceptual processing of acoustic rhythms. Thus, we predict that motor synchronization with auditory rhythms containing simple or complex frequency ratios should yield greater power of neural oscillations and more negative amplitudes within a typical N1 time window at the frequencies with a simple ratio than those with the complex ratio.

We investigated the neural correlates of auditory–motor rhythm processing during perception, production, and synchronization of auditory rhythms containing simple and complex frequency ratios. Two primary questions were addressed: The first question was whether neural entrainment with a simple-ratio auditory rhythm is enhanced compared to complex-ratio rhythms, in perception, production, and synchronization tasks. Neural entrainment was measured by assessing neural oscillations in the frequency domain, and ERPs time-locked to auditory stimulus onsets and to motor responses in the time domain. The second question was whether behavioral synchronization is enhanced in response to simple rhythms compared to complex rhythms and whether behavioral synchronization is associated with neural entrainment measures. We examined accuracy and stability of synchronization in a coordination task in which individuals tapped along with auditory rhythms that formed simple and complex ratios.

The current study tested perception and production of auditory rhythms by skilled musicians, who are experienced

with the perception and production of both simple- and complex-ratio rhythms and therefore offer an ideal population for studying rhythmic behavior (Collier & Wright, 1995). Participants first listened to auditory sequences comprising two-part rhythms (Listen task), each part presented with a different constant pitch. One part occurred at a constant (fixed) frequency across Rhythm conditions, whereas the other part varied in frequency across conditions relative to the fixed frequency, to form a 1:1 integer ratio (1:1 condition), a 1:2 integer ratio (1:2 condition), or a 3:2 integer ratio (3:2 condition). Rhythmic complexity therefore ranged from low (1:1) to moderate (1:2) to high (3:2), consistent with predictions of Farey tree hierarchies (Glass & Mackey, 1988). Participants then performed a Synchronize task in which they tapped at the fixed frequency while aiming to synchronize their movements with the other part. Stimulus-to-tap ratios varied across three Rhythm conditions with ratios of 1:1, 1:2, and 3:2, consistent with the rhythmic ratios in the Listen task. To achieve a baseline measure of cortical responses during rhythmic movement in the absence of synchronization, participants completed a control Motor task, in which they tapped at the same fixed frequency as during the Synchronization task. Neural activity was recorded at the scalp using EEG during the Listen, Synchronize, and Motor tasks, and sound corresponding to the stimulus and/or taps was present in all conditions.

Following predictions from nonlinear dynamical systems, neural oscillations should exhibit most power in the simple-ratio (1:1) Rhythm condition and least power in the most complex (3:2) Rhythm condition for both Listen and Synchronize conditions. Furthermore, the amplitude of neural oscillations at the constant frequency, as well as the amplitude of ERP waveforms within an N1-related time window, should increase during the Synchronization task compared to the Listen task. Participants in the Synchronize task should exhibit greatest synchrony of tapping with the auditory stimulus in the simple-ratio (1:1) Rhythm condition and least synchrony in the complex-ratio (3:2) Rhythm condition. Behavioral synchrony measures and the amplitude of ERP waveforms within the N1 time window are expected to decrease together as rhythmic complexity increases (from 1:1 to 1:2 to 3:2). Comparison of N1 amplitudes across tasks (Listen/Synchronize/Motor) and Rhythm conditions (1:1, 1:2, 3:2) should reveal how interactions between auditory perceptual and motor processes give rise to accurate auditory-motor synchronization. We examine these interactions under naturalistic stimulus conditions such as those that occur during music and speech production, in which auditory feedback from both stimuli and responses are present.

## METHODS

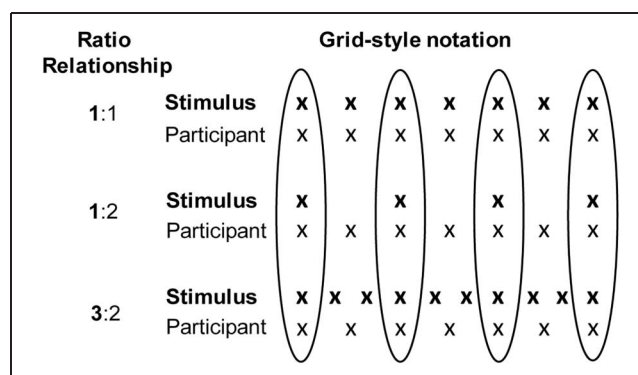
### Participants

Twenty-nine adults (21 women, eight men; aged 18–30 years;  $M = 22.6$  years,  $SD = 3.1$  years) with at least 6 years of

formal instruction on a musical instrument (range: 6–17 years;  $M = 9.2$  years,  $SD = 2.9$  years) participated. Participants currently practiced their instrument an average of 5.8 hr a week ( $SD = 8.6$ ) and averaged 12.2 years of experience playing their instrument ( $SD = 4.4$ ). All participants were right-handed and did not possess any neurological disorders. Participants passed an audiometric screening test, in which they demonstrated hearing thresholds  $\leq 20$  dB for 1000-, 750-, 500-, and 250-Hz tones (representing the range of pitches used in the experiment). Three additional participants were recruited. Two of these participants' data were excluded because of poor EEG signal quality, and a third participant's data were excluded because of low task performance (hit rates that exceeded 3 SDs from the group mean in the Listen task). The study was reviewed by the McGill University research ethics board. Participants provided a written consent to participate after they were fully informed about the study.

### Stimuli

The auditory stimuli were sequences of repeating rhythmic patterns (see Figure 1). The sequences consisted of low-pitched (392-Hz) sine tones and high-pitched (660-Hz) woodblock sounds, which were perceptually distinct in pitch, timbre, and presentation rate. Three different Rhythm complexity conditions were created from the high- and low-pitched tone sequences: a 1:1 ratio, a 1:2 ratio, and a 3:2 ratio (the first number indicates the rate of the high pitch, and the second number indicates the rate of the low pitch). The low-pitched tones were presented at a constant rate with an interonset interval (IOI) of 528 msec, whereas the IOI for the high-pitched tones differed across the rate ratios with 528 msec for the 1:1, 1056 msec for the 1:2, and 352 msec for the 3:2 condition. Thus, the stimulus ("high-pitched part") frequency was defined as 1.89 Hz in the 1:1 condition, 0.94 Hz in the 1:2 condition, and 2.84 Hz in the 3:2 condition. The prescribed tap ("low-pitched part") frequency was 1.89 Hz across all Rhythm conditions and tasks. The shared frequency (which corresponded to



**Figure 1.** Schematic showing the alignment of participant tap onsets (indicated by non-bold Xs) and stimulus onsets (indicated by bold Xs) across a cycle of tap onsets for each of three Rhythm conditions. Circles indicate the shared stimulus frequency across Rhythm conditions.

points of simultaneity between the two frequencies across all Rhythm conditions, shown by circles in Figure 1) was defined as 0.94 Hz. The timbres were generated on a sound module SD-50 SoundCanvas (Roland Inc.) using “Pre\_333 Simple sine” (Synth Lead category) for the low-pitched tone and “Rhy 001” (Drums category) for the high-pitched tone. The attack time of the sine tone and woodblock sounds was 5 msec, and the decay time for both sounds was 36 msec. The same percussion sound was used for a metronome cue that signaled the tapping tempo at the start of each trial. The sound pressure level of high- and low-pitched tones and metronome tones was set to 75 dB SPL, confirmed with a sound level meter.

Each trial contained four metronome clicks with IOIs of 528 msec, followed for 30 sec by one of the three Rhythms (1:1, 1:2, and 3:2). Twelve trials were created for each of the three Rhythm conditions in the Listen task. Two of the 12 trials contained one missing beat in either the low-pitched part or the high-pitched part. Twelve trials were created for each of the three Rhythm conditions in the Synchronize task; each trial contained the four-beat metronome cue followed by the high-pitched part (without missing beats), during which participants heard the low-pitched part when they tapped. Twelve trials were included in the Motor task; each trial presented the four-beat metronome cue, after which participants heard the low-pitched part when they tapped.

## Equipment

The experiment took place in a sound-attenuated and electrically shielded testing room. The audiometric screening was administered with a diagnostic audiometer using over-ear headphones provided by Maico (MA-40, Maico GmbH). Auditory stimuli were presented over EEG-compatible insert earphones (ER1-14B, Etymotic Research). Participants tapped to the low-pitched sequence by pressing a key (note name C3) on an electronic keyboard (PSR-500M, Yamaha Inc.) that transmitted timing information with 1-msec resolution via a musical instrument digital interface (Yamaha Inc.). Information about the tap timing was recorded using FTAP software (Finney, 2001) modified to transmit event triggers (Mathias, Gehring, & Palmer, 2017).

Participants wore an EEG cap with 64 Ag/AgCl electrodes configured according to an extension of the International 10–20 system. EEG signals were recorded by a BioSemi ActiveTwo system at a resolution of 24 bits and a sampling rate of 1024 Hz (BioSemi, Inc.). The EEG was grounded using BioSemi’s combination of common mode sense and drive right leg electrodes. Electrodes below and above the right eye monitored vertical eye movements, and two electrodes placed adjacent to the outer canthi of the eyes monitored horizontal eye movements.

## Design

The within-participant  $2 \times 3$  repeated-measures design included two factors: Task (Listen, Synchronize) and

Rhythm condition (1:1, 1:2, 3:2). The Motor condition served as a separate control to allow identification of a motor ROI for analysis of power spectral density (PSD). Each participant completed the tasks in this order—Listen, Synchronize, and Motor condition—to ensure that tapping rates did not influence the perceptual neural responses in the Listen task. The order of tasks was fixed to ensure that Listen blocks were not influenced by prior experience with producing the stimulus rhythms (via auditory or motor imagery; cf. Brown & Palmer, 2013). Within Listen and Synchronize tasks, the blocks of Rhythm condition trials were presented in a fixed order of 1:1, 1:2, and 3:2, ranging from the easiest to the most difficult. As practice effects should favor the final (most complex rhythm) condition within each block (Tajima & Chosi, 2000), the order of Rhythm conditions was fixed to bias away from the hypothesis that synchronization performance should be best for simple relative to complex rhythmic ratios.

There were two practice trials and 12 experimental trials for each Rhythm condition (1:1, 1:2, 3:2) within the Listen and Synchronize tasks and two practice trials and 12 experimental trials for the Motor task, yielding 84 experimental trials ( $12 \times 3 \times 2 + 12 = 84$ ).

## Procedure

Participants first provided written consent, completed a questionnaire about their musical training background, and completed an audiometry screening. Participants who were not able to detect any tones presented at or below 20 dB were excluded from the experiment. Participants were then outfitted with an EEG cap and electrodes and completed the experimental tasks.

### *Listen Task*

Participants first listened separately to the sine tone and the woodblock sequences to become familiarized with the auditory stimuli. The woodblock sequence was referred to by the experimenter as the “high-pitched part,” and the sine tone sequence was referred to as the “low-pitched part.” The participants then heard a sample of a Listen trial containing both parts for each Rhythm condition (1:1, 1:2, and 3:2 ratios). Participants were instructed to listen and report, at the end of the trial, any missing sounds in either part (following Nozaradan, Zerouali, Peretz, & Mouraux, 2015). Each Rhythm condition began with two 38.54-sec practice trials, one that contained an “omitted” beat and one with a “no-omitted” beat; participants were offered more practice trials if they desired. The participants completed 12 test trials for each Rhythm condition, of which two contained missing beats, as well as the two practice trials. Test trials were 38.5 sec each; therefore, each block (practice and test trials) comprised approximately 9 min ( $38.5 \text{ sec} \times 2 \text{ practice trials} + 38.5 \text{ sec} \times 12 \text{ test trials}$ ) of testing plus short breaks between successive trials.

### Synchronization Task

Participants were first familiarized with the two rhythmic parts in each sequence by listening to them separately and then together for each rhythm condition. They then completed two 38.5-sec practice trials, followed by the 12 test trials for each rhythm condition. Participants were instructed to listen to four metronome clicks and then, after the metronome stopped, to start tapping with their right (dominant) hand at the rate of the metronome while synchronizing with the presented tone sequence (synchronization–continuation). They were told that their goal was to synchronize their taps with the high-pitched part such that their taps formed the specified rhythmic ratio (1:1, 1:2, or 3:2 depending on the condition) with the high-pitched part. Participants completed two practice trials and 12 experimental trials for each of the 1:1, 1:2, and 3:2 conditions.

### Motor Task

Participants were presented with the four-beat metronome cue at the same tapping rate (528-msec IOI) at which they tapped across all conditions. They were asked to tap at the rate presented by the metronome cue during the 38.5-sec trial until their taps no longer produced sound, which signaled the end of the trial. Participants completed two practice trials and 12 experimental trials. After this condition, participants removed the EEG cap and received a small compensation.

The entire experiment lasted about 2 hr. During the experiment, participants were monitored by an experimenter who invited participants to take breaks between trials and blocks and who offered water between blocks.

### Behavioral Data Analysis

Hit rates (percentage of trials with correct detection of missing tones) and false alarm rates (percentage of trials with incorrect detection of missing tones) were computed for each Rhythm condition in the Listen task. Intertap intervals (ITIs) between consecutive taps were computed for the Synchronize and Motor tasks as the difference between each pair of adjacent tap onsets, and mean ITIs were computed by averaging ITIs within an analysis window that excluded the first and last four beats of each trial, leaving 60 taps per trial. Coefficients of variation (CVs) of ITIs were computed for each trial as the standard deviation divided by the mean ITI. Absolute asynchrony of participant taps with the auditory stimulus was computed as the absolute difference between each tap onset and the most temporally proximal stimulus onset ( $|\text{tap onset} - \text{stimulus onset}|$ ). Signed asynchrony was computed as the tap onset minus the most temporally proximal stimulus onset; thus, a negative value indicates anticipatory behavior. No participants were classified as outliers (3 *SDs* or

more from the group mean) in terms of their mean ITIs or mean absolute asynchronies.

### EEG Data Analysis

The EEG data were preprocessed in the EEGLAB software package (Delorme & Makeig, 2004). Raw continuous EEG data were first down-sampled to 512 Hz (*pop\_resample.m*) and referenced to the common average across electrode sites (*pop\_reref.m*). Independent component analysis (ICA) was subsequently used to identify and remove stereotypical eye blink and lateral eye movement artifacts from the data analysis (Jung et al., 2000; Bell & Sejnowski, 1995).

ICA was computed on a version of the original data that was preprocessed using procedures previously shown to optimize ICA component identification (Debener, Thorne, Schneider, & Viola, 2010; code for those procedures has been published in Stropahl, Bauer, Debener, & Bleichner, 2018); the version of the data submitted to ICA is hereafter referred to as the ICA set. First, the ICA set was filtered with application of low-pass (40 Hz, Order 100) and high-pass (1 Hz, Order 500) Hanning windowed sinc finite impulse response filters (*pop\_firws.m*). Bad channels were then visually identified and removed to reduce noise contributions to ICA decomposition. Filtered data were then parsed into short (1-sec) epochs to identify transient nonstereotypical artifacts (such as sudden body movements), which are typically short-lasting and nonperiodic. Any 1-sec epochs greater than 2 *SDs* from the mean activity across segments and channels were identified (*pop\_jointprob.m*) and removed from the data. Cleaned data epochs were then submitted to infomax ICA (*pop\_runica.m*), which reconstructs continuous time series from epoched data. To account for possible loss of rank from common average referencing, the option “PCA” was used in the ICA algorithm to set the number of decomposed components to equal one less than the number of channels. Stereotypical eye artifacts (eye blinks and lateral eye movements) for each participant were visually identified from their respective ICA set. ICA weights for each participant were subsequently applied to their original continuous data sets to remove identified eye artifacts (*pop\_subcomp.m*); the same bad channels that were removed from ICA sets were removed from original sets before application of ICA weights to ensure consistency of dimensions.

The application of ICA weights to original data sets resulted in artifact-attenuated sets that were subsequently preprocessed using a different set of procedures tailored for planned time- and frequency-domain analyses. First, data in bad channels were spherically interpolated from neighboring channels (*pop\_interp.m*). Second, additional noise was removed through the application of low-pass (20 Hz, Order 1000) and high-pass (0.1 Hz, Order 1000) Hanning windowed sinc finite impulse response filters tailored for the planned time- and frequency-domain analyses (Zamm et al., 2017).

Trials in the Listen condition with incorrect responses regarding tone omissions were identified and excluded from all planned EEG analyses.

### PSD

Artifact-corrected EEG data were assessed for spectral content using a procedure adapted from Zamm et al. (2017). Continuous data in each trial were segmented into 10.56-sec consecutive (nonoverlapping) epochs, corresponding to 20 taps per epoch at 528 msec per tap. Epoch edges were multiplied with a 5407-sample (10.56-sec) Hanning window, corresponding to the duration of each epoch. Data segmentation allowed for the exclusion of epochs in trials in which participants responded incorrectly regarding tone omissions in the Listen task.

Each epoch was subsequently submitted to PSD estimation (*pwelch.m* in MATLAB), using a window length equivalent to the epoch duration (5407 samples) and no overlap specified (overlap = []); this implementation uses no subwindow and is therefore equivalent to Bartlett's method (Bartlett, 1950). The PSD was estimated at the stimulus frequency (1.89, 0.94, and 2.84 Hz in 1:1, 1:2, and 3:2 conditions, respectively), the tap frequency (1.89 Hz across all conditions), and the shared frequency (0.94 Hz across all conditions), as described earlier. The resulting power spectra were log-transformed ( $10 \cdot \log_{10}$ , dB conversion) and subsequently averaged across trials for each channel and every participant. Similar to other EEG studies (Zamm et al., 2017; Tierney & Kraus, 2014; Nozaradan et al., 2011, 2012), a noise reduction procedure was then applied to ensure reduced influence of residual spectral noise on each channel, by subtracting from each frequency the mean power at  $\pm 3$  neighboring frequency bins, corresponding to  $\pm 0.1875$  Hz. The noise reduction outcomes could yield a flat spectrum centered around 0 if the signal contained only noise or a peak resulting from the noise-subtracted spectrum if the signal contained nonnoise components. Immediately adjacent frequencies were included in the noise estimates to capture a point-by-point estimate of spectral change (Zamm et al., 2017).

Noise-subtracted spectra were averaged across all channels for each participant (Tierney & Kraus, 2014; Nozaradan et al., 2011, 2012), and PSDs at the target frequencies of 1.89, 0.94, and 2.84 Hz were extracted from each participant's power spectrum and exported for subsequent analyses. Auditory and motor ROIs were defined by the electrodes that displayed maximal PSD across participants in grand-averaged topographies for Listen (averaged across the three Rhythm conditions) and Motor tasks respectively, following Nozaradan et al. (2012); these ROIs were electrode FCz in the Listen task and electrode C3 in the Motor task. Both electrodes (FCz and C3) were evaluated as ROI in the Synchronize task. None of the sensors within these ROIs included interpolated channels.

### ERPs

Both tap- and stimulus-locked ERPs were computed. EEG data were segmented into 600-msec epochs time-locked to the participants' taps and to the auditory stimuli with a 100-msec baseline period. Epochs time-locked to missing tones in the Listen condition were excluded from the analyses. Average ERP waveforms were computed for each participant for the Listen and Synchronize conditions. Amplitudes within a stereotypical N1 time window were statistically evaluated at electrodes Fz and FCz at a latency of 80–120 msec after previous research on auditory–motor tasks (Mathias, Gehring, & Palmer, 2019; Mathias et al., 2017; Horváth & Burgyán, 2013; Barry, 2009; Katahira, Abla, Masuda, & Okanoya, 2008). The Fz and FCz sensors did not correspond to interpolated channels for any participant.

## RESULTS

### Behavioral Results

#### Listen Task

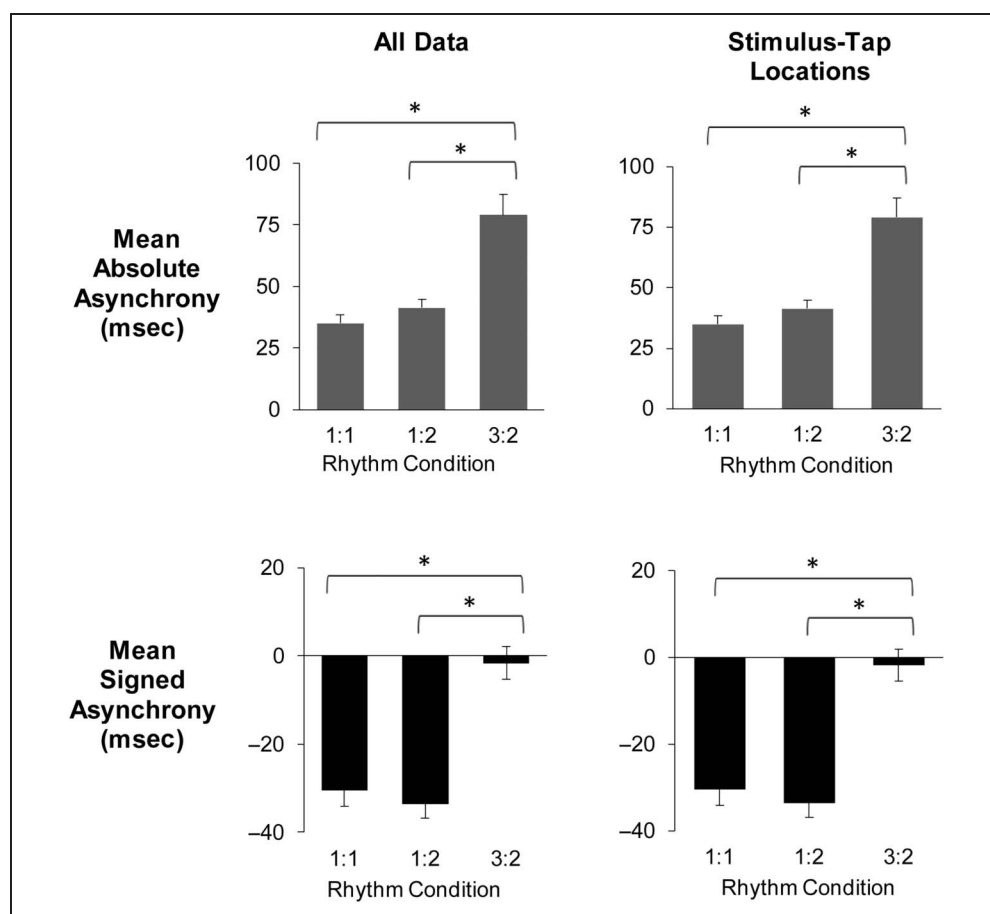
Successful detection of missing sounds in the Listen trials was measured by the hit rate (percentage of trials with correct detection). Responses indicating a missing sound when one did not occur were measured by the false alarm rate (percentage of trials with incorrect detection). The mean false alarm rate was 1.4% in the 1:1 Rhythm condition, 6.2% in the 1:2 condition, and 4.1% in the 3:2 condition. Because false alarm rates were very low in all conditions (less than one trial per condition or 10%), the analysis of Rhythm condition effects focused on hit rates. A one-way repeated-measures ANOVA on mean hit rates by Rhythm condition yielded a significant main effect,  $F(1, 28) = 1128.89$ ,  $p < .001$ . The mean hit rate was 96.6% in the 1:1 condition ( $SE = 2.4\%$ ), 91.4% in the 1:2 condition ( $SE = 3.6\%$ ), and 86.2% in the 3:2 condition ( $SE = 4.2\%$ ). Tukey post hoc tests revealed that the mean hit rate was larger in the 1:1 condition than in the 3:2 condition (Tukey HSD = 9.46,  $\alpha = .05$ ). Thus, participants were least accurate at detecting the missing sound onsets in the complex 3:2 Rhythm condition.

#### Synchronize Task

Mean ITIs were 528 msec ( $SE = 0.4$  msec) in the 1:1 condition, 529 msec ( $SE = 0.4$  msec) in the 1:2 condition, and 530 msec ( $SE = 2.7$  msec) in the 3:2 condition. The ITIs did not significantly differ between conditions ( $p > .05$ ). There was a significant effect of Rhythm condition on the mean CV,  $F(2, 56) = 8.18$ ,  $p < .001$ . The mean CV was larger in the 3:2 Rhythm condition (mean CV = 0.075) than in the 1:1 condition (mean CV = 0.05) and the 1:2 condition (mean CV = 0.055; Tukey HSD = 0.0195,  $\alpha = .01$ ).

Absolute asynchronies between tap and stimulus onsets in the Synchronize trials were computed for each Rhythm condition. There was a significant effect of Rhythm

**Figure 2.** Mean absolute asynchrony of participants' taps (top) and mean signed asynchronies (tap onset minus stimulus onset; bottom) in the Synchronization task by Rhythm condition. Analyses based on all tap events are shown in the left column; analyses of the subset of taps at the shared frequency (taps that aligned with stimulus onsets across Rhythm conditions, circled in Figure 1) are shown in the right column. \* $p < .01$ . Error bars represent 1  $SE$ .



condition on absolute asynchrony values,  $F(2, 56) = 30.08$ ,  $p < .001$ . As indicated in Figure 2, participants were more asynchronous in the 3:2 condition than in the 1:1 and 1:2 conditions (Tukey HSD = 18.59,  $\alpha = .01$ ). Thus, participants showed greater variability as well as reduced synchronization accuracy in the 3:2 Rhythm condition.

As shown in Figure 1, the subset of participants' taps that aligned temporally with stimulus onsets varied across the 1:1, 1:2, and 3:2 Rhythm conditions. To control for differences in number of asynchrony values among rhythm conditions, we re-analyzed the subset of asynchronies for taps that aligned with stimulus onsets. The same analysis repeated on these absolute asynchronies confirmed the main effect of Rhythm condition,  $F(2, 56) = 30.16$ ,  $p < .001$ . Participants were more asynchronous in the 3:2 Rhythm condition than in the 1:1 and 1:2 conditions (Tukey HSD = 18.58,  $\alpha = .01$ ). The standard deviations of absolute asynchronies for the same subset of taps were also re-analyzed; the ANOVA yielded the same main effect of rhythm condition,  $F(2, 56) = 52.11$ ,  $p < .001$ . Participants' asynchronies were more variable in the 3:2 Rhythm condition than in the 1:1 and 1:2 conditions (Tukey HSD = 9.96,  $\alpha = .01$ ). Thus, analysis of asynchronies that controlled for the number of stimulus-tap events yielded the same results as the analysis of all participant taps.

Signed asynchronies (tap onset - stimulus onset) in the Synchronize trials were computed for each Rhythm

condition. There was a significant effect of Rhythm condition on signed asynchrony values,  $F(2, 56) = 32.07$ ,  $p < .001$ . As indicated in Figure 2, participants' signed asynchronies were significantly more anticipatory in the 1:1 and 1:2 conditions than in the 3:2 condition (Tukey HSD = 12.32,  $\alpha = .01$ ). A reanalysis of the subset of asynchronies for the taps that aligned with stimulus onsets that are shared across the Rhythm conditions (circled in Figure 1) confirmed the main effect of Rhythm condition,  $F(2, 56) = 37.03$ ,  $p < .001$ . Participants' signed asynchronies were significantly more negative in the 1:1 and 1:2 conditions than in the 3:2 condition (Tukey HSD = 12.34,  $\alpha = .01$ ). Thus, participants' taps showed greater anticipation of the auditory stimulus in the 1:1 and 1:2 Rhythm conditions than in the 3:2 Rhythm condition.

### Motor Task

Participants' mean ITI in the Motor task was 514 msec ( $SE = 3$  msec), slightly shorter than the prescribed interval of 528 msec. A regression analysis predicting the mean ITI by the serial position within each trial ( $n = 62$  ITIs) revealed that the participants sped up and shortened the tapping interval by 0.12 msec per ITI ( $r = -.76$ ,  $p < .01$ ). The mean CV of ITIs was 0.062 ( $SE = 0.0067$ ), similar to the mean CV in Synchronize trials ( $M = 0.060$ ). These results served as a control for the accuracy and

precision of participants' tapping when an external auditory stimulus is absent.

In summary, both perceptual detection of missing stimulus onsets and synchronization were enhanced in response to simple multivoiced rhythms compared to complex rhythms.

## EEG Results

### PSD

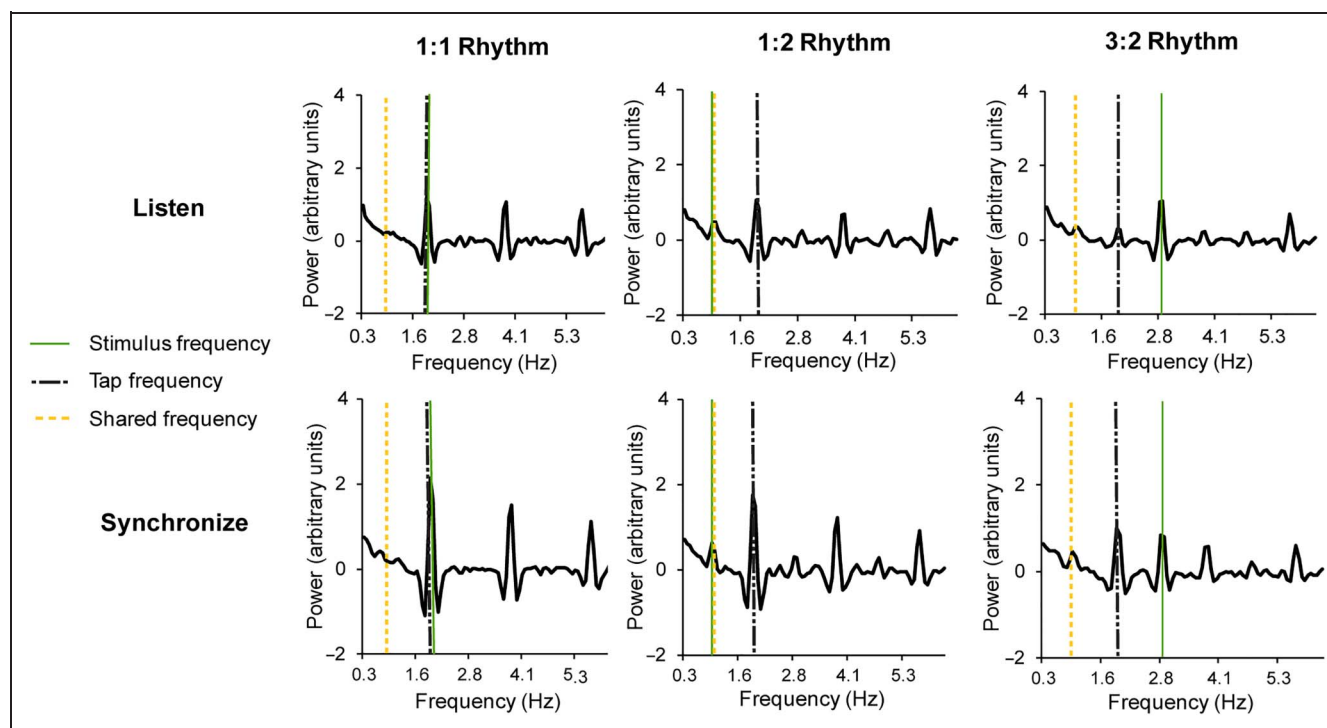
Figure 3 shows the mean power spectra for each task and Rhythm condition averaged across electrodes. Prominent peaks occurred at or near frequencies corresponding to the stimulus rates (0.94, 1.89, and 2.84 Hz) and multiples of the stimulus rate. Topographic maps of the peak PSD for each task, Rhythm condition, and frequency of interest are shown in Figure 4, which indicate characteristic patterns of auditory cortex activity with maximum at frontal midline electrodes (FCz) in the Listen task, sensorimotor activity at left central electrodes (C3) in the Motor task, and both activities in the Synchronize task.

*Stimulus frequency effects.* A three-way ANOVA was conducted on the mean spectral power at the stimulus frequency with the factors Task (Listen/Synchronize), Rhythm condition, and ROI (auditory/motor). We compared the frequency of the stimulus voice that participants tapped along with (did not produce) in the Synchronize

task with the corresponding stimulus frequency in the Listen task.

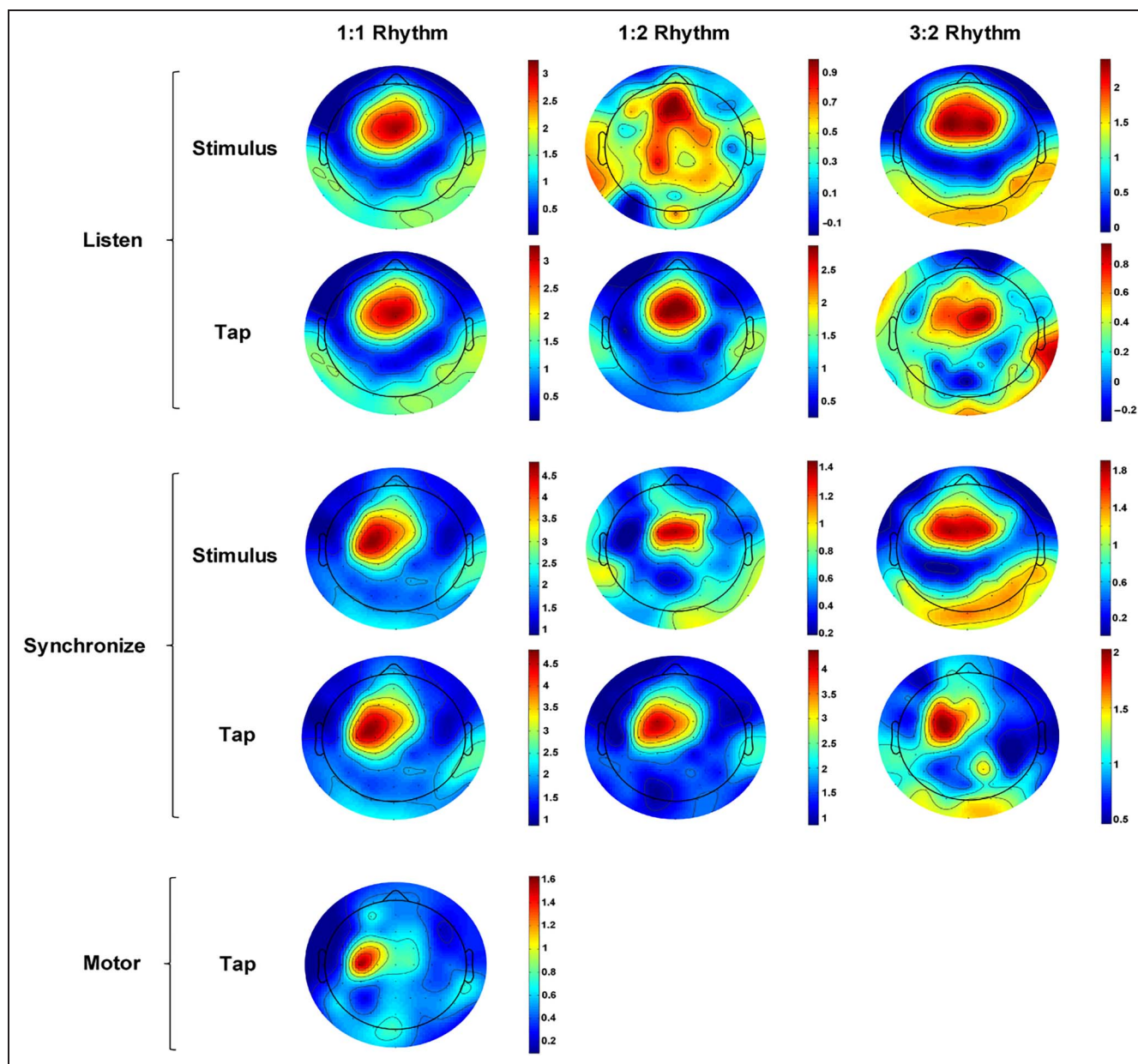
The ANOVA revealed a main effect of Task,  $F(1, 28) = 29.35, p < .001$ . Spectral power at the stimulus frequency was greater during the Synchronize than during the Listen tasks. The ANOVA also revealed main effects of Rhythm condition,  $F(2, 56) = 74.53, p < .001$ , and ROI,  $F(1, 28) = 31.54, p < .001$ . Mean spectral power was significantly greater at the 1:1 Rhythm condition than at the 3:2 condition, and the 3:2 condition was significantly greater than the 1:2 condition (Tukey HSD = 0.49,  $p < .05$ ). There was also a significant Task  $\times$  Rhythm condition interaction,  $F(2, 56) = 23.93, p < .001$ , and significant Task  $\times$  Rhythm condition  $\times$  ROI interaction,  $F(2, 56) = 13.67, p < .001$ . To pursue the interactions, two-way ANOVAs were conducted on spectral power at the stimulus frequency for each ROI separately.

The mean spectral power present at the stimulus frequency in the auditory ROI differed significantly by Task,  $F(1, 28) = 6.38, p < .05$ , and by Rhythm condition,  $F(2, 56) = 55.92, p < .001$ , with greater power during Synchronize than Listen tasks, and for the 1:1 Rhythm than for other Rhythm conditions (Tukey HSD = 0.74,  $p < .01$ ). The interaction of Task and Rhythm condition was also significant,  $F(2, 56) = 6.96, p < .01$ . As shown in Figure 5, the power for the Listen task was greatest in 1:1, followed by 3:2, and least in 1:2 Rhythms; power in the Synchronize task was greater for 1:1 than for other rhythms (Tukey HSD = 0.82,  $p < .05$ ).



**Figure 3.** Mean power spectra for the Listen task (top) and the Synchronize task (bottom) by Rhythm condition (1:1, 1:2, and 3:2 ratios). Stimulus frequencies for each condition are shown with a continuous vertical line, the tapping frequency (1.89 Hz) is shown with a dotted vertical line, and the shared frequency (0.94 Hz) is indicated with a dashed vertical line.





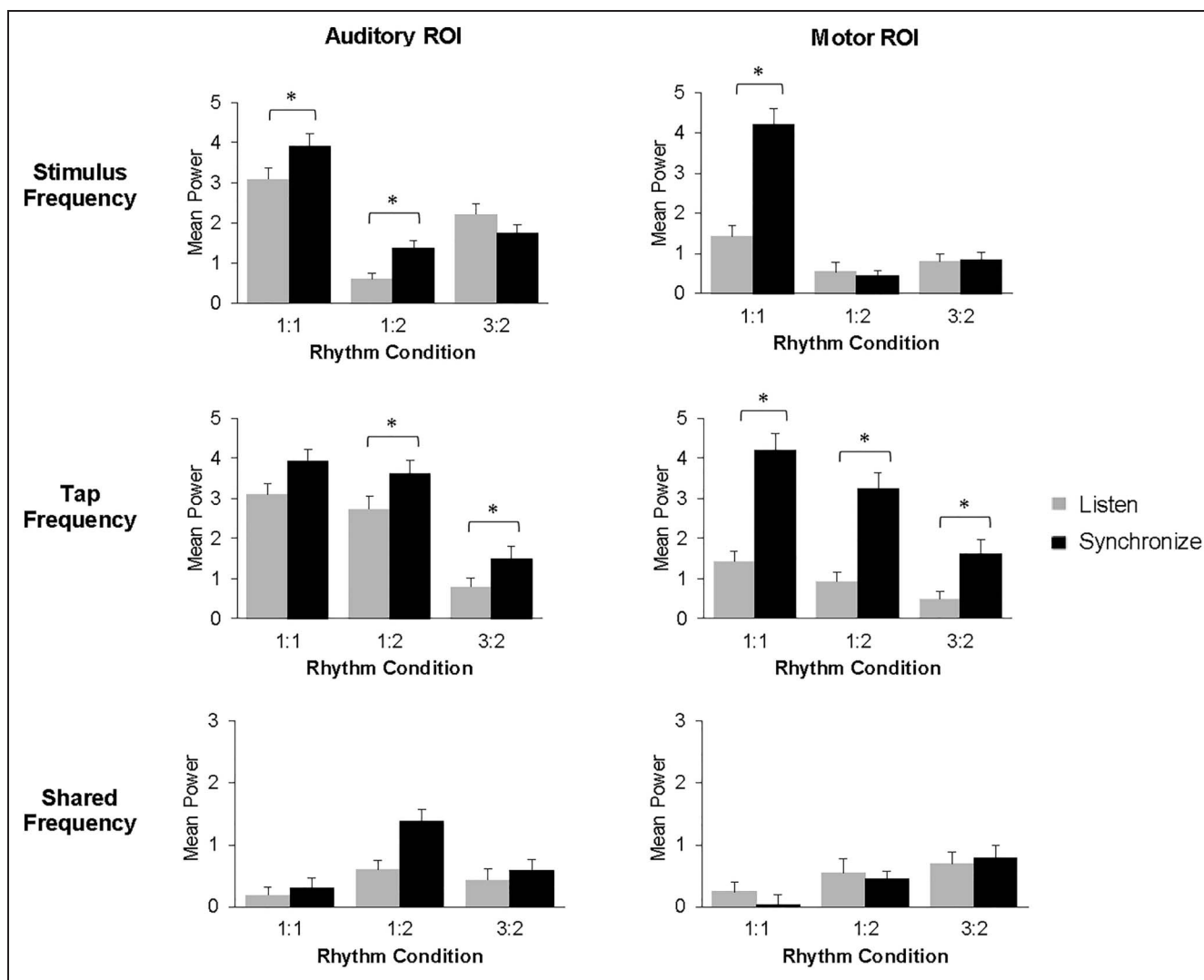
**Figure 4.** Mean PSD scalp topographies at the stimulus frequencies (1.89 Hz in the 1:1 Rhythm condition, 0.94 Hz in the 1:2 Rhythm condition, and 2.84 Hz in the 3:2 Rhythm condition), Tap frequency (1.89 Hz) by Rhythm condition. Top: Listen task. Center: Synchronize task. Bottom: Motor task. Scalp topographies are scaled within condition, ranging from the minimum PSD to the maximum PSD.

The same two-way ANOVA on the spectral power present in the motor ROI also indicated main effects of Task,  $F(1, 28) = 29.23, p < .001$ , and of Rhythm condition,  $F(2, 56) = 49.01, p < .001$ . Mean spectral power was greater for Synchronize than for Listen tasks and for 1:1 Rhythm conditions than for 1:2 or 3:2 Rhythms (Tukey HSD = 0.77,  $p < .01$ ). There was a significant interaction of Task with Rhythm condition,  $F(2, 56) = 30.38, p < .001$ . As shown in Figure 5, the 1:1 Rhythm condition in the Listen task yielded more power than the 1:2 and 3:2 conditions (Tukey HSD = 1.04,  $p < .01$ ). The 1:1 Rhythm condition in the Synchronize task also yielded significantly greater power than all other rhythms in the Synchronize task and

all Rhythm conditions in the Listen task (Tukey HSD = 1.04,  $p < .01$ ).

In summary, effects of the stimulus frequency indicated that the motor ROI exhibited greater mean power during the Synchronize task with the 1:1 Rhythm, compared to the other Rhythm conditions and to the Listen task. As expected, the auditory ROI indicated greater power for the 1:1 Rhythm compared to other rhythms in both the Listen and Synchronize tasks.

**Tap frequency effects.** The same ANOVA on mean spectral power at the tap frequency revealed a main effect of Task,  $F(1, 28) = 44.19, p < .001$ . Spectral power at the tap



**Figure 5.** Mean spectral power at the stimulus frequencies (top), tap frequency (middle), and shared frequency (bottom), by Task, ROI, and Rhythm condition. The left column shows mean power within the auditory ROI, and the right column shows mean power within the motor ROI. Gray bars represent the Listen task, and black bars represent the Synchronize task. Significant pairwise comparisons are shown between means that differed only for the stimulus and tap frequencies, as those frequencies showed a significant three-way Task  $\times$  Rhythm  $\times$  ROI interaction. \* $p < .05$ . Error bars represent 1 SE.

frequency was greater during the Synchronize condition than during the Listen condition, as expected. There was also a main effect of Rhythm condition,  $F(2, 56) = 65.49$ ,  $p < .001$ , and ROI,  $F(1, 28) = 8.06$ ,  $p < .01$ . Similar to findings at the stimulus frequencies, mean spectral power was greatest at the 1:1 Rhythm condition; however, power at the tap frequency was greater in the 3:2 condition than in the 1:2 condition (Tukey HSD = 0.45,  $p < .05$ ). There were significant Task  $\times$  Rhythm condition interaction,  $F(2, 56) = 4.99$ ,  $p = .01$ , Task  $\times$  ROI interaction,  $F(1, 28) = 15.40$ ,  $p = .001$ , Rhythm condition  $\times$  ROI interaction,  $F(2, 56) = 8.13$ ,  $p = .001$ , and three-way Task  $\times$  Rhythm condition  $\times$  ROI interaction,  $F(2, 56) = 3.94$ ,  $p < .05$ . Two-way ANOVAs were conducted at each ROI to address the complex interactions.

The mean spectral power measured at the auditory ROI for the tap frequency indicated significant effects of both

tasks,  $F(1, 28) = 14.33$ ,  $p < .01$ , and of Rhythm condition,  $F(2, 56) = 60.97$ ,  $p < .001$ . As expected, spectral power at the tap frequency was greater in Synchronize than in Listen conditions; power was greater in the 1:1 and 1:2 conditions compared to the 3:2 condition (Tukey HSD = 0.70,  $p < .01$ ). There was no significant interaction.

The mean spectral power measured at the motor ROI for the tap frequency also yielded significant main effects of Task,  $F(1, 28) = 43.02$ ,  $p < .001$ , and of Rhythm condition,  $F(2, 56) = 32.96$ ,  $p < .001$ . In addition, the interaction of Task and Rhythm condition was significant,  $F(2, 56) = 10.40$ ,  $p < .001$ . As shown in Figure 5, spectral power at the motor ROI was greater for Synchronize tasks than for Listen tasks and greater for the 1:1 Rhythm, followed by the 1:2 Rhythm and the 3:2 Rhythm (Tukey HSD = 0.66,  $p < .01$ ). In addition, spectral power was greater in the Synchronize task for all three Rhythm conditions than

in the Listen task, with the largest difference between tasks in the 1:1 Rhythm condition (Tukey HSD = 0.94,  $p < .01$ ).

In summary, analyses of spectral power at the tap frequency showed similar findings to analyses at the stimulus frequency. Spectral power was greater in response to the simple 1:1 Rhythm than the more complex rhythms. Both auditory and motor ROIs showed greater power for the Synchronize task compared with the Listen task.

**Shared frequency results.** The same three-way ANOVA conducted on spectral power at the shared frequency (0.94 Hz) across Rhythm conditions (circled in Figure 1) revealed a main effect of Rhythm condition,  $F(2, 56) = 8.78, p < .001$ . Spectral power was significantly greater for the 1:2 and 3:2 Rhythm conditions compared to the 1:1 condition (Tukey HSD = 0.42,  $p < .001$ ). There were also significant Task  $\times$  ROI interaction,  $F(1, 28) = 1.54, p < .05$ , and Rhythm condition  $\times$  ROI interaction,  $F(2, 56) = 4.78, p < .05$ . Two-way ANOVAs were conducted on each ROI to pursue the complex interactions.

The two-way ANOVA on spectral power in the auditory ROI at the shared frequency indicated significant main effects of Task,  $F(1, 28) = 4.82, p < .05$ , and Rhythm condition,  $F(2, 56) = 9.96, p < .001$ ; there were no significant interactions. As shown in Figure 5, spectral power was greater in Synchronize tasks than in Listen tasks. In contrast to stimulus frequency and tap frequency findings, spectral power at the shared frequency was greatest in the 1:2 Rhythm condition and less in the 1:1 and 3:2 conditions (Tukey HSD = 0.38,  $p < .05$ ).

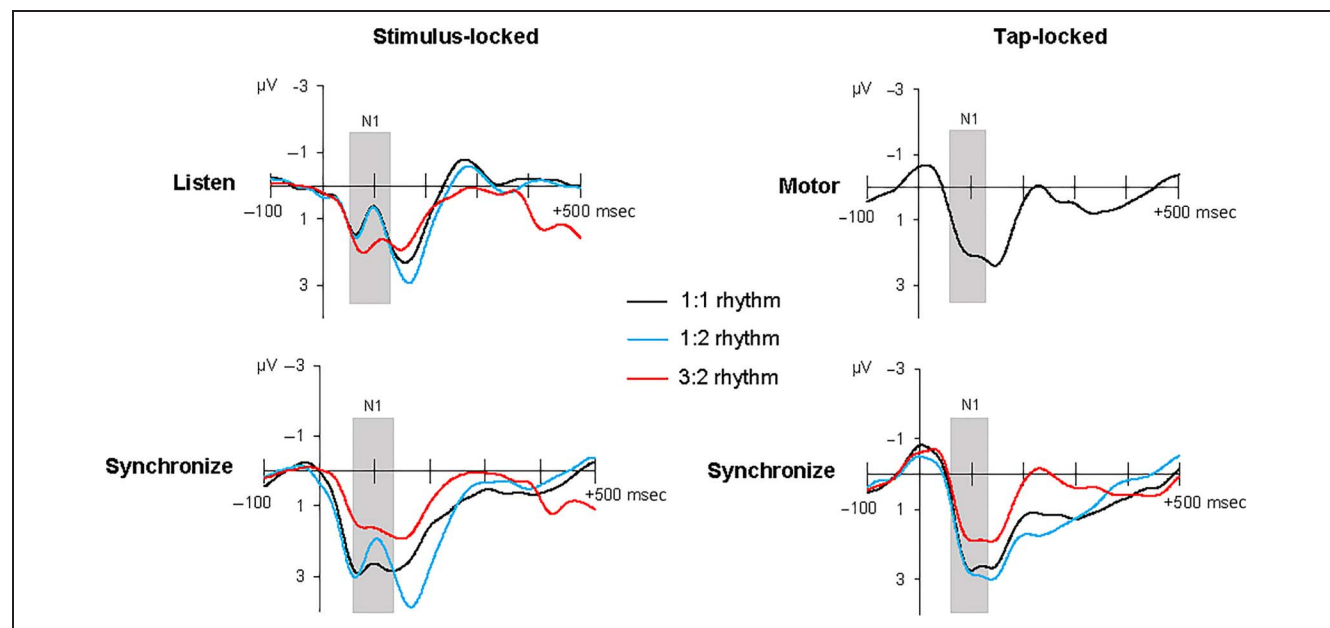
The two-way ANOVA on spectral power in the motor ROI at the shared frequency indicated a significant main

effect of Rhythm condition,  $F(2, 56) = 4.94, p < .05$ . Spectral power was significantly greater in the 3:2 Rhythm condition than in the 1:1 Rhythm condition (Tukey HSD = 0.59,  $p < .01$ ). There were no significant main effects of Task or interaction. Thus, in contrast to the auditory ROI findings of greatest power at the shared frequency for the 1:2 Rhythm, the motor ROI indicated increased power for the 3:2 (most difficult) condition.

### ERPs

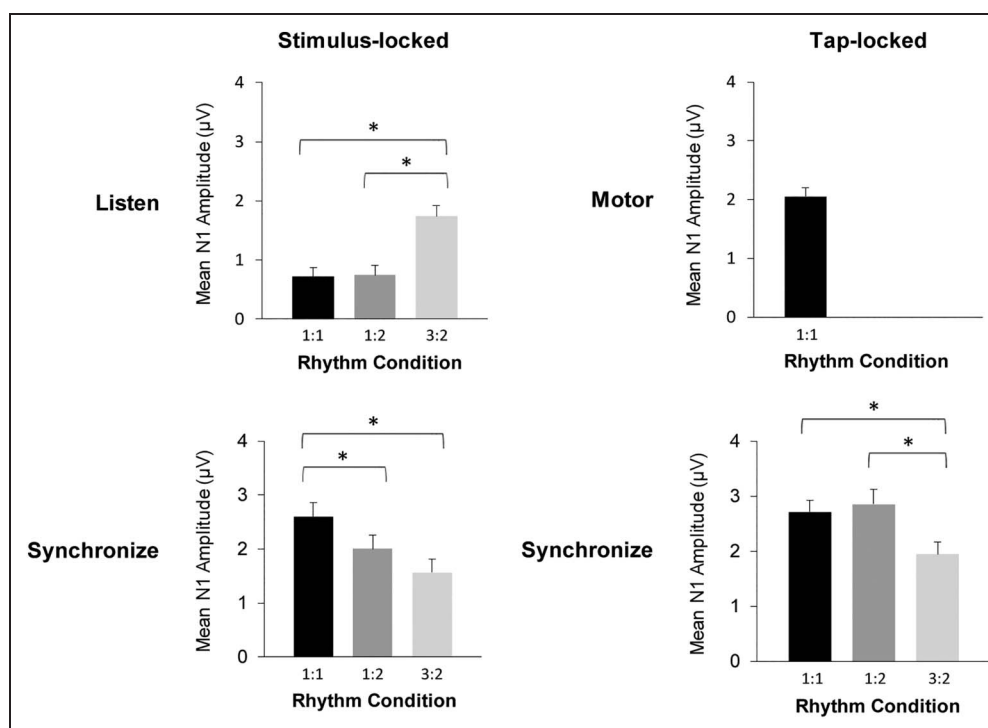
Effects of rhythmic complexity on the amplitude of ERP waveforms were examined in the Listen and Synchronize tasks. Figure 6 shows the grand-averaged ERP waveforms time-locked to tap onsets and to stimulus onsets. To control for potential differences in the number of stimulus tones between Rhythm conditions, we analyzed event-related responses elicited by taps and stimuli for only the subset of locations at which taps and stimuli aligned across Rhythm conditions (circled in Figure 1). This ensured that the same number of events was included across Rhythm conditions as well as in stimulus-locked and tap-locked analyses in the Synchronization condition. The onset times for stimulus-locked ERPs elicited by high- and low-pitched stimuli (the two parts of the auditory stimuli) were identical in the Listen condition; onset times for stimulus- and tap-locked ERPs were not identical for the Synchronize condition, because participants did not always tap synchronously with the stimulus.

**Listen task.** We first assessed ERP amplitudes in the N1 time window across Rhythm conditions in the Listen task, shown in Figure 7. A one-way ANOVA on stimulus-locked



**Figure 6.** Grand-averaged ERPs elicited during the Listen (top left), Motor (top right), and Synchronize (bottom) tasks by Rhythm condition (1:1, 1:2, and 3:2) at electrode Fz. ERPs time-locked to the stimulus onsets are shown in the left column. ERPs time-locked to participant taps are shown in the right column. Shaded areas indicate the N1 ERP component. Negative values are plotted upward.

**Figure 7.** Mean N1 amplitude values in the Listen (top left), Motor (top right), and Synchronize (bottom) tasks by Rhythm condition (1:1, 1:2, and 3:2). Stimulus-locked N1 amplitudes are shown in the left column, and tap-locked N1 amplitudes are shown in the right column. Positive values are plotted upward.  $*p < .05$ . Error bars represent 1 SE.



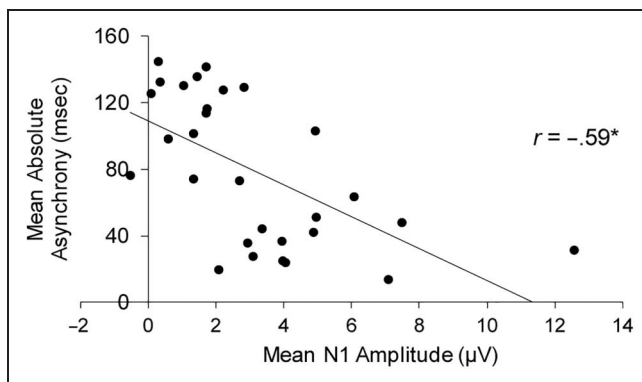
mean amplitudes in the N1 time window yielded a main effect of Rhythm,  $F(2, 56) = 24.53, p < .001$ . The stimulus tones in the 3:2 Rhythm condition elicited less negative N1 amplitudes than in both the 1:1 and 1:2 Rhythm conditions (Tukey HSD = 0.50,  $p < .01$ ). Thus, participants demonstrated a reduction in mean amplitudes within the N1 time window while listening to the 3:2 Rhythm compared to the 1:1 and 1:2 Rhythms.

**Synchronization task.** We assessed mean amplitudes in the N1 time window across rhythms in the Synchronize task (also shown in Figure 7). A one-way ANOVA on stimulus-locked amplitudes yielded a main effect of Rhythm,  $F(2, 56) = 13.11, p < .001$ . Both the 3:2 and 1:2 Rhythm conditions elicited more negative mean amplitudes than the 1:1 condition (Tukey HSD = 0.49,  $p < .05$ ). The tap-locked amplitudes in the synchronization task also yielded a main effect of Rhythm,  $F(2, 56) = 10.08, p < .001$ . The 3:2 Rhythm condition elicited more negative amplitudes in the N1 time window than both the 1:1 and 1:2 conditions (Tukey HSD = 0.66,  $p < .01$ ). Thus, amplitudes in the N1 time window that were time-locked to both taps and to stimuli were more negative for the 3:2 Rhythm condition than the 1:1 condition.

**Motor task.** We compared mean amplitudes in the N1 time window elicited during the Motor task with mean amplitudes observed in the Listen and Synchronize tasks. A one-way ANOVA conducted on mean amplitudes for the 1:1 Rhythm (tap-locked N1 amplitudes in the Motor and Synchronize tasks and stimulus-locked amplitudes in the Listen task) yielded a main effect of Task,  $F(2, 56) = 56.8,$

$p < .001$ . Amplitudes in the N1 time window were more suppressed (more positive) in the Motor and Synchronize conditions compared to the Listen condition and less suppressed (more negative) in the Motor condition compared to the Synchronize condition (Tukey HSD = 0.58,  $p < .01$ ).

**Correlations between behavioral asynchrony, N1 amplitudes, and individual differences in musical practice.** We tested the relationship between participants' asynchronies, amplitudes in the N1 time window, and amount of weekly musical practice (which ranged from 0 to 30 hr/week) in the most complex synchronization condition, the 3:2 Rhythm condition. A multiple linear regression was conducted to predict absolute asynchronies in the 3:2 condition from participants' hours of weekly instrumental practice, stimulus- and tap-locked ERP amplitudes within the N1 time window in the 3:2 condition, and PSD measures at the stimulus and tap frequencies in the 3:2 condition. A significant regression was observed,  $R^2 = .74, F(5, 23) = 5.50, p < .005$ . Semi-partial correlations indicated significant contributions to asynchrony from stimulus-locked amplitudes,  $\beta = -.73, t(23) = 2.76, p < .05$ , and from hours of weekly practice,  $\beta = -.38, t(23) = 2.39, p < .05$ . As stimulus-locked amplitudes in the N1 time window became more positive, absolute asynchrony in the 3:2 Rhythm condition decreased, and as weekly practice increased, absolute asynchrony in the 3:2 Rhythm condition decreased. The two significant predictors did not correlate with each other,  $r(27) = -.12, p = .53$ . Thus, musical practice and stimulus-locked amplitudes in the N1 time window yielded independent contributions to absolute asynchrony in the 3:2 Synchronize condition.



**Figure 8.** Correlation between mean stimulus-locked N1 amplitudes and mean absolute asynchronies in the 3:2 Rhythm condition of the Synchronize task.  $*p < .05$ .

Figure 8 shows the simple correlation of mean absolute asynchronies in the 3:2 Rhythm condition with mean stimulus-locked amplitudes in the N1 time window,  $r(27) = -.59, p < .05$ . Simple correlations of the absolute asynchronies with mean amplitudes in the N1 time window and with hours of weekly musical practice did not reach significance in the 1:1 (stimulus-locked:  $r(27) = -.09, p > .05$ ; musical practice:  $r(27) = -.33, p = .08$ ) or 1:2 (stimulus-locked:  $r(27) = .003$ ; musical practice:  $r(27) = -.28; ps > .05$ ) Rhythm conditions. No significant simple correlations between hours of weekly musical practice and stimulus-locked amplitudes in the N1 time window were observed. Moreover, there were no significant correlations between amplitudes in the N1 time window and PSD amplitudes.

In summary, multiple regression analyses revealed that more negative amplitudes in the N1 time window and greater amounts of weekly musical practice were independently associated with greater synchronization accuracy for complex rhythms.

## DISCUSSION

We examined behavioral and neural entrainment during musicians' naturalistic perception and production of simple and complex auditory rhythms. Participants listened to two-part auditory sequences whose rhythms formed integer ratios varying in complexity from low (1:1) to moderate (1:2) and high (3:2) complexity. One of the two parts was fixed in rate across all rhythmic complexity conditions, allowing us to compare neural responses to the two parts under similar conditions. Participants also performed a synchronization task with the same rhythms in which they tapped at the fixed rate while synchronizing with the other auditory part. Finally, participants performed a motor task in which they tapped the same fixed rate in the absence of other auditory stimuli. In contrast to many studies of auditory-motor synchronization, participants' taps resulted in auditory feedback in all conditions of the current study. As auditory feedback typically

accompanies movement in natural synchronization tasks such as music performance (for a review, see Palmer, 2013), the current design provides a step forward in identifying mechanisms of auditory-motor coordination under more natural feedback conditions comparable to musicians' performance.

The musicians performed at high levels of overall accuracy in both perceptual and production tasks. Nonetheless, their behavioral responses indicated effects of rhythmic complexity in both perceptual and production tasks. Detection rates for missing beats in the Listen task worsened in the most complex (3:2) rhythm condition. Participants were most variable in tapping durations and least synchronous with stimulus onsets in the 3:2 rhythm condition in the Synchronize task, consistent with previous findings that behavioral entrainment decreases as rhythmic complexity increases (Chapin et al., 2010; Collier & Wright, 1995). Finally, participants' tapping accuracy remained high in the Motor task (in the absence of other auditory stimuli). Importantly, participants' accuracy of tapping the fixed frequency was equivalently high across Synchronize and Motor tasks, allowing comparison of neural entrainment in the presence of equivalent behavior. Several studies document musicians' greater accuracy and precision in producing rhythms (Summers, Rosenbaum, Burns, & Ford, 1993; Summers & Kennedy, 1992) and perceiving rhythms (Manning & Schutz, 2016). Thus, musicians' rhythmic behavior, often near ceiling, provides a conservative test of the hypothesis that rhythmic complexity modulates synchronization performance.

Neural measures of entrainment to the stimulus and tap frequencies also decreased as rhythmic complexity increased during Listen and Synchronize tasks. Specifically, the 1:1 rhythm condition elicited greater entrainment (higher PSD) at stimulus and tap frequencies relative to other rhythm conditions (1:1, 1:2), at both auditory and motor ROIs. This finding supports dynamical systems models of rhythmic entrainment (Large et al., 2015; Strogatz, 2001), specifically Farey tree frameworks of rhythmic stability, which suggest that rhythms featuring simple integer ratios should display more stable entrainment relative to rhythms featuring complex integer ratios (Bouvet, Varlet, Dalla Bella, Keller, & Bardy, 2017; Peper, Beek, & Van Wieringen, 1991). The observed relationship between neural entrainment and rhythmic complexity is consistent with our behavioral results, which also indicated enhanced entrainment for simple relative to complex rhythms. Future research could address how rhythmic complexity modulates other characteristics of entrainment than period coupling of neural oscillations with an external stimulus. For example, entrainment is also characterized by the dynamic alignment of oscillatory phase with a stimulus (Bauer, Bleichner, Jaeger, Thorne, & Debener, 2018); the dynamics of phase alignment may vary as a function of rhythmic complexity, whereby phase alignment occurs more rapidly with simple relative to complex rhythms.

Neural entrainment was also modulated by task: Enhanced EEG spectral density was observed at stimulus and tap frequencies during synchronization relative to perception (Listen task), for both auditory and motor ROIs. This finding is consistent with previous studies showing enhanced amplitudes of neural oscillations at musical beat frequencies in perceived and produced tone sequences (Fujioka et al., 2015; Nozaradan et al., 2011, 2012). Moreover, the current findings extend this work by revealing consistent effects of synchronization on entrainment across simple and complex rhythms.

We also assessed neural entrainment to the frequency at which stimulus and taps aligned (referred to as the shared frequency; see Figure 1). This frequency corresponded to either the stimulus or tap frequency in the 1:1 and 1:2 conditions (stimulus and tap frequencies were integer multiples) but did not correspond to either stimulus or tap frequency in the 3:2 condition. Most importantly, in the 3:2 condition, the shared frequency emerged from the poly-rhythmic relationship between stimulus and tap frequencies. Enhanced auditory neural entrainment to the shared frequency occurred in the Synchronize relative to the Listen task, suggesting that tapping along with a stimulus amplifies one's perception of the frequency at which individuals align taps with a rhythmic stimulus. Importantly, this effect cannot be explained by the increased acoustic amplitude at the shared frequency arising from temporally coincident stimulus and tap onsets, as this was controlled across rhythm conditions and tasks.

We also investigated how rhythmic complexity modulated event-related cortical responses to tone onsets, specifically within the time window of the N1 ERP component. The N1 component has been linked to enhanced auditory perceptual processing (Nobre & van Ede, 2018; Lange et al., 2003; Näätänen & Winkler, 1999), as would be expected for attended frequencies in perceived and produced rhythms. Increases in rhythmic complexity in both Listen and Synchronize tasks resulted in more positive amplitudes in the N1 time window in the 3:2 rhythm condition relative to the 1:1 and 1:2 conditions. Modulation of the N1 time window by rhythmic complexity contributes to a growing literature that specifies the mechanisms underlying the occurrence and stability of spontaneous movement synchronization to auditory rhythms (Bouvet et al., 2019).

Changes in ERP amplitudes within the N1 time window were also modulated by whether participants perceived or produced rhythms. Amplitudes in the Motor and Synchronize task for the simplest rhythm condition (1:1) were more positive relative to the Listen task. In addition, amplitudes in the Motor task were more negative relative to the Synchronize task. These findings are consistent with the view that N1 changes reflect motor-induced suppression of auditory cortical processing (Horváth, 2015; SanMiguel, Todd, & Schröger, 2013). They also fit with evidence that the N1 wave of auditory evoked responses is attenuated in response to sound that is not selectively attended to,

relative to attended sound (Snyder, Alain, & Picton, 2006; Hillyard et al., 1973). The current findings further indicate that task complexity modulates the N1; we propose that the Synchronize task (which presented one auditory part to track) required a higher level of auditory processing than was required in the Motor task (with no parts to track) and a lower level of auditory processing than was required in the Listen task (with two parts to track).

The observed pattern of amplitudes within the N1 time window may also be influenced by other mechanisms. It is possible that overlap of ERP responses to temporally proximal tone onsets may have resulted in waveform cancellation in the N1 analysis time window (carryover effects). For example, the P1 elicited by a given tap onset could have overlapped in time with the N1 response to a preceding stimulus tone, which may have resulted in voltage cancellation and potentially reduced amplitudes within the N1 time window. Amplitudes in stereotypical N1 time windows may also be influenced by refractory effects from stimulus rates. It is well established that tones presented in shorter intervals elicit N1 ERPs of a smaller magnitude than tones presented at longer intervals (Budd et al., 1998), whereas in contrast, the P1 and P2 waves are less affected by increases in stimulus rate (Gutschalk, Patterson, Uppenkamp, Scherg, & Rupp, 2004). Decreases in N1 ERP magnitude at faster stimulus rates may arise from decreased excitability of neural generators at intervals smaller than the refractory period of the underlying network (Gutschalk et al., 2004). The observed pattern of results cannot be fully accounted for by any one of these mechanisms; future work should aim to disentangle how each mechanism may differentially contribute to ERP responses elicited by rhythms of varying complexity. This could potentially be accomplished using alternative measures of ERP amplitude such as peak-to-peak approaches (Snyder et al., 2006), which may disambiguate neighboring ERPs within an analysis time window.

Finally, musicians' behavioral asynchrony measures in the high-complexity (3:2) Synchronize task decreased as amplitudes within the N1 time window became more positive and the amount of musical practice increased. The more positive amplitudes were associated with smaller asynchronies in the 3:2 rhythm condition, suggesting that motor-induced suppression of auditory cortical processing aided synchronization; this effect was larger in participants with greater amounts of musical training. Although amplitudes within the N1 time window can be sensitive to learning effects, decreasing as listeners adapt to stimulus repetition over multiple blocks (Ross, Barat, & Fujioka, 2017), the Synchronization condition in the current study occurred after the Listen condition, by which time participants had become familiar with the auditory stimuli from all rhythmic complexity conditions. Thus, learning-related adaptation is less likely to account for the decreased amplitudes within the N1 time window observed during the Synchronize condition. Another experimental ordering consideration is the role of imagery; a music production task followed by a

perceptual task that relies on similar stimulus material can lead to the use of imagery (auditory or motor) during the later perceptual task (Mathias, Palmer, Perrin, & Tillmann, 2015; Brown & Palmer, 2013). To avoid these potential imagery effects, the current study ordered the Listen task before the Synchronize and Motor tasks, and participants were not informed about the Synchronize and Motor tasks until after the Listen task had been completed. Future studies may manipulate the order of perception and production tasks in the presence of auditory feedback to further evaluate learning effects on amplitudes within the N1 time window.

In summary, the current study compared behavioral and neural responses across perception and production tasks and levels of rhythm complexity while controlling for the participants' tapping rate and while providing auditory feedback associated with both stimuli and responses. These controls allowed us to compare directly the neural responses to simple and complex rhythms presented in comparable naturalistic conditions in terms of perceptual stimulation and response rates. Although previous studies have attempted to control auditory feedback across perception conditions (Nozaradan et al., 2011, 2012) or response rates across motor conditions (Mathias et al., 2017), to our knowledge, no single study has simultaneously controlled motor production rate and the presence of self-generated auditory feedback. This study has demonstrated that behavioral and neural entrainment underlies accurate auditory-motor synchronization and is modulated in similar ways by rhythmic complexity. Many real-world auditory synchronization tasks—such as group music performance—contain actions that are accompanied by auditory feedback; the current study represents a step toward understanding more naturalistic sensorimotor synchronization behaviors. Future studies may address how rhythmic complexity modulates entrainment across individuals in multiperson synchronization tasks (such as group music-making) as well as how nonexperts (such as musical novices) learn to entrain to complex auditory rhythms.

## Acknowledgments

This research was funded in part by an NSF Graduate Fellowship to B. Mathias, a PBEEE Graduate award from FRQNT to A. Zamm, an NSERC-USRA award to P. Gianferrara, and NSERC Grant 298173 and a Canada Research Chair to C. Palmer. We thank Shelby Trapid, James O'Callaghan, Jamie Dunkle, and Frances Spidle for assistance.

Reprint requests should be sent to Caroline Palmer, Department of Psychology, McGill University, Montreal, Quebec, Canada H3A 1B1, or via e-mail: caroline.palmer@mcgill.ca.

## REFERENCES

- Barry, R. J. (2009). Evoked activity and EEG phase resetting in the genesis of auditory Go/NoGo ERPs. *Biological Psychology*, *80*, 292–299.
- Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, *37*, 1–16.
- Bauer, A. R., Bleichner, M. G., Jaeger, M., Thorne, J. D., & Debener, S. (2018). Dynamic phase alignment of ongoing auditory cortex oscillations. *Neuroimage*, *167*, 396–407.
- Bays, P. M., & Wolpert, D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *Journal of Physiology*, *578*, 387–396.
- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129–1159.
- Bood, R. J., Nijssen, M., Van Der Kamp, J., & Roerdink, M. (2013). The power of auditory-motor synchronization in sports: Enhancing running performance by coupling cadence with the right beats. *PLoS One*, *8*, e70758.
- Bouvet, C. J., Varlet, M., Dalla Bella, S., Keller, P. E., & Bardy, B. G. (2017). Auditory-motor entrainment to complex frequency ratios. In J. A. West-Knapp & G. J. Pepping (Eds.), *Studies in Perception and Action XIV: Nineteenth International Conference on Perception and Action* (pp. 45–48). New York: Routledge.
- Bouvet, C. J., Varlet, M., Dalla Bella, S., Keller, P. E., Zelic, G., & Bardy, B. G. (2019). Preferred frequency ratios for spontaneous auditory-motor synchronization: Dynamical stability and hysteresis. *Acta Psychologica*, *196*, 33–41.
- Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, *7*, 320.
- Brown, S., & Parsons, L. M. (2008). The neuroscience of dance. *Scientific American*, *299*, 78–83.
- Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs. refractoriness. *International Journal of Psychophysiology*, *31*, 51–68.
- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*, 1926–1929.
- Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S. J., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: The role of attending. *Frontiers in Psychology*, *1*, 224.
- Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body movement selectively shapes the neural representation of musical rhythms. *Psychological Science*, *25*, 2147–2159.
- Collier, G. L., & Wright, C. E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 602–627.
- Debener, S., Thorne, J. D., Schneider, T. R., & Viola, F. C. (2010). Using ICA for the analysis of multi-channel EEG Data. In M. Ullsperger & S. Debener (Eds.), *Simultaneous EEG and fMRI: Recording, analysis, and application* (pp. 121–133). New York: Oxford Scholarship Online.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2*, 704–716.
- Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, *33*, 65–72.
- Fitzroy, A. B., & Sanders, L. D. (2015). Musical meter modulates the allocation of attention across time. *Journal of Cognitive Neuroscience*, *27*, 2339–2351.
- Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, *35*, 15187–15198.
- Glass, L., & Mackey, M. C. (1988). *From clocks to chaos: The rhythms of life*. Princeton, NJ: Princeton University Press.

- Gutschalk, A., Patterson, R. D., Uppenkamp, S., Scherg, M., & Rupp, A. (2004). Recovery and refractoriness of auditory evoked fields after gaps in click trains. *European Journal of Neuroscience*, *20*, 3141–3147.
- Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, *51*, 347–356.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65.
- Horváth, J., & Burgyán, A. (2013). No evidence for peripheral mechanism attenuating auditory ERPs to self-induced tones. *Psychophysiology*, *50*, 563–569.
- Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, *24*, 1919–1931.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T., McKeown, M. J., Iragui, I., et al. (2000). Removing electroencephalographic artefacts by blind source separation. *Psychophysiology*, *37*, 163–178.
- Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: An ERP study. *Neuroscience Research*, *61*, 120–128.
- Kelso, J. A. S. (1991). Multifrequency behavioural patterns and the phase attractive circle map. *Biological Cybernetics*, *64*, 485–495.
- Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event-related potential study. *Psychophysiology*, *40*, 806–817.
- Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, *9*, 159.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*, 119–159.
- Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.
- Manning, F. C., & Schutz, M. (2016). Trained to keep a beat: Movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychological Research*, *80*, 532–542.
- Mathias, B., Gehring, W. J., & Palmer, C. (2017). Auditory N1 reveals planning and monitoring processes during music performance. *Psychophysiology*, *54*, 235–247.
- Mathias, B., Gehring, W. J., & Palmer, C. (2019). Electrical brain responses reveal sequential constraints on planning during music performance. *Brain Sciences*, *9*, 25.
- Mathias, B., Palmer, C., Perrin, F., & Tillmann, B. (2015). Sensorimotor learning enhances expectations during auditory perception. *Cerebral Cortex*, *15*, 2238–2254.
- Miura, A., Kudo, K., & Nakazawa, K. (2013). Action–perception coordination dynamics of whole-body rhythmic movement in stance: A comparison study of street dancers and non-dancers. *Neuroscience Letters*, *544*, 157–162.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*, 826–859.
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, *19*, 34–48.
- Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual differences in rhythmic cortical entrainment correlate with predictive behaviour in sensorimotor synchronization. *Scientific Reports*, *6*, 20612.
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*, 10234–10240.
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, *32*, 17572–17581.
- Nozaradan, S., Schönwiesner, M., Caron-Desrochers, L., & Lehmann, A. (2016). Enhanced brainstem and cortical encoding of sound during synchronized movement. *NeuroImage*, *142*, 231–240.
- Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex*, *25*, 736–747.
- Palmer, C. (2013). Music performance: Movement and coordination. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 405–422). Amsterdam: Elsevier Press.
- Peper, C. E., Beek, P. J., & Van Wieringen, P. C. W. (1991). Bifurcations in polyrhythmic tapping: In search of Farey principles. In J. Requin & G. Stelmach (Eds.), *Tutorials in motor neuroscience* (pp. 413–431). Dordrecht, The Netherlands: Springer.
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*, 403–452.
- Ross, B., Barat, M., & Fujioka, T. (2017). Sound-making actions lead to immediate plastic changes of neuromagnetic evoked responses and induced  $\beta$ -band oscillations during perception. *Journal of Neuroscience*, *37*, 5948–5959.
- SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, *50*, 334–343.
- SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*, 8633–8639.
- Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research*, *75*, 95–106.
- Schroeder, M. (1991). *Fractals, chaos, power laws*. New York: Freeman.
- Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*, 1–13.
- Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*, 117–126.
- Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental Brain Research*, *222*, 149–157.
- Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*, 268–276.
- Stropahl, M., Bauer, A.-K., Debener, S., & Bleichner, M. G. (2018). Source-modeling auditory processes of EEG data using EEGLAB and Brainstorm. *Frontiers in Neuroscience*, *12*, 309.
- Summers, J. J., & Kennedy, T. M. (1992). Strategies in the production of a 5:3 polyrhythm. *Human Movement Science*, *11*, 101–112.
- Summers, J. J., Rosenbaum, D., Burns, B., & Ford, S. (1993). Production of polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 416–428.
- Tajima, M., & Choshi, K. (2000). Effects of learning and movement frequency on polyrhythmic tapping performance. *Perceptual and Motor Skills*, *90*, 675–690.



- Thaut, M. (2013). *Rhythm, music, and the brain: Scientific foundations and clinical applications*. New York: Routledge.
- Tierney, A., & Kraus, N. (2014). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, 27, 400–408.
- Wing, A. M., Endo, S., Bradbury, A., & Vorberg, D. (2014). Optimal feedback correction in string quartet synchronization. *Journal of the Royal Society Interface*, 11, 20131125.
- Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, 72, 2031–2046.
- Zamm, A., Palmer, C., Bauer, A.-K. R., Bleichner, M. G., Demos, A. P., & Debener, S. (2017). Synchronizing MIDI and wireless EEG measurements during natural piano performance. *Brain Research*, 1716, 27–38.