Brain and Cognition 136 (2019) 103594

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Ternary meter from spatial sounds: Differences in neural entrainment between musicians and non-musicians

Alexandre Celma-Miralles[a,*], Juan M. Toro[a,b]

[a] Universitat Pompeu Fabra, C. Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain
[b] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluis Companys, 23, 08019 Barcelona, Spain

ARTICLE INFO

Keywords:
Beat perception
Meter induction
Spatial audition
EEG
Frequency-tagging

ABSTRACT

The present study explores the relationship between the rhythmic structure of music and the spatial dimension of sound. We study how the brain interacts with spatially-separated sounds to build up a metrical structure. Participants listened to sequences of isochronous sounds that came from different positions on the azimuth plane: 0° (control condition), ± 30°, ± 60° or ± 90° (spatial conditions). Ternary meter was signaled by the alternation of one sound on one side and two sounds on the symmetrical side. In Experiment 1, musicians and non-musicians paid attention to the spatial sounds. In Experiment 2, participants paid attention to a visual distractor. We recorded their electroencephalograms and performed frequency-tagging analyses. In both experiments, the isochronous beat elicited steady-state evoked-potentials at the frequency of the beat (2.4 Hz). While in Experiment 1 the alternation produced clear responses at the frequency of the ternary meter (0.8 Hz), in Experiment 2 these responses were only significant in the Spatial 90° condition, and mainly in musicians. This suggests that top-down attentional mechanisms are in play for meter induction. Besides, musicians showed stronger responses to beat and meter than non-musicians, suggesting that formal musical training enhances the neural entrainment to spatially-defined rhythms.

1. Introduction

Music is the aesthetical arrangement of tones over time. The concatenation of musical tones in rhythms creates melodies that are governed by overlapping metrical and tonal-harmonic structures. Our brain processes the rhythms underlying musical excerpts in a very organized way (Fitch, 2013). First, an isochronous beat is extracted from the stream of sounds, creating the perception of periodic points over time. Secondly, these periodic sounds are organized hierarchically into sequences of strong and weak patterns. These two processes are generally termed as beat perception and meter induction (Fitch, 2013). Depending on the specific arrangement of strong and weak events, different metrical structures arise in music and dance. The strong beat (i.e. downbeat) occurs at an integer multiple (a subharmonic frequency) of the beat, usually at integer ratios such as 2:1 or 3:1. For instance, the *strong-weak* combination produces the binary meter (like a march), while the *strong-weak-weak* arrangement leads to the ternary meter (like a waltz). This organization of periodic rhythmic events into metrical structures seems to be a biologically-rooted human universal (Ravignani, Delgado, & Kirby, 2016; Ravignani, Thompson, Grossi, Delgado, & Kirby, 2018; Savage, Brown, Sakai, & Currie, 2015).

Most of the times, the saliency of the downbeat depends on physical variations of the sounds. For instance, the downbeat of a musical piece is normally marked by changes in loudness, pitch or timbre (London, 2012). However, music is not a purely acoustic event, and it might also involve information from other sensory modalities. This may include vision and the spatial domain, as in the synchronized movements of dancing, conducting and group performing (Bishop & Goebl, 2017; Lee, Barrett, Kim, Lim, & Lee, 2015; Luck & Sloboda, 2007). In dance, for instance, the beat and its rhythmic structure may be signaled by changes in posture, movement velocity and spatial position (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2014; Toiviainen, Luck, & Thompson, 2010). In orchestra conducting, the speed and the amplitude of the gesture also signals the tempo of the beat and its metrical structure. The involvement of these other modalities in the processing of musical features makes interesting to explore the role of space in our rhythmic perception.

In music, the production of sounds at distinct locations allows composers to create movement effects and textures that alter the perception of the listeners. The use of sound localization as a musical resource is nowadays termed as "spatialization". Although the term for spatial music was coined recently in the Western music tradition,

* Corresponding author.
E-mail address: alexandre.celma@upf.edu (A. Celma-Miralles).

https://doi.org/10.1016/j.bandc.2019.103594
Received 4 September 2018; Received in revised form 2 August 2019; Accepted 3 August 2019
Available online 12 August 2019

masses and multiple-choir pieces from the 15th century, such as those composed by Giovanni Pierluigi da Palestrina and Alessandro Striggio, already exploited this dimension (Moroney, 2007). However, it was not until the development of music technology during the last century that spatialization became more relevant (Born, 2013; Harley, 2016). Furthermore, the invention of accurate music reproduction devices that deal with multiple dispositions of sound sources boosted the research on the spatial dimension of sound, pushing sound technicians to improve and develop better systems for contemporary music and cinema (Pachet & Delerue, 1998). In this context, it is important to discover if spatial cues can modulate the neural processing of the metrical structure of music.

The neural entrainment to beat and meter has been studied indirectly, analyzing the EEG responses to violations of expected metrical structures (Winkler, Denham, & Nelken, 2009), and directly, observing the EEG recordings of rhythmic excerpts in the frequency dimension (Nozaradan, 2014, 2017). Here, we will analyze the fluctuations of the neural populations that follow the beat and the meter in the frequency domain. The ongoing cortical activity elicited during the presentation of periodic stimuli is known as steady-state evoked-potentials. Several fMRI studies reported the involvement of motor areas in rhythmic processing (Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Schubotz, Friederici, & Von Cramon, 2000), so it is also our hypothesis that regions such as the premotor cortex, supplementary motor areas, basal ganglia and cerebellum might allow for neural entrainment to metrical structures over different domains, including the acoustic (Nozaradan, 2014) and the visual domain (Celma-Miralles, de Menezes, & Toro, 2016). In these last two studies, top-down processes were recruited to establish the metrical structure of the beat across auditory and visual domains. In the present study we will test whether the domain generality of meter induction extends to the spatial domain as well.

The present work explores the neural bases of meter induction over auditory spatial cues. In the following experiments, we focus on whether the spatial distribution of sounds might allow the brain to synchronize to the ternary meter. We selected this metrical structure because Fujioka, Zendel, and Ross (2010) found more contrasts in the auditory evoked responses between strong and weak beats for the ternary meter than for the binary meter. Our participants repetitively listened to an alternation of three sounds at two symmetrical spatial locations: the first sound came from one side and the two following sounds came from the contralateral side (see Fig. 1). We controlled the spatial location of the tones we presented to the participants by modifying two binaural cues: the Interaural Time Differences (ITD) and the Interaural Level Differences (ILD). While other cues are informative in the vertical plane, such as the filtering effects of each individual pinna, auditory channel and head position (Butler & Belendiuk, 1977), the modulation of ITD and ILD are essential for the localization of sounds on the horizontal plane (Gilkey & Anderson, 2014). By modifying binaural cues (see Fig. 2), we have the opportunity to explore the neural entrainment to the ternary meter extracted from the distribution of sounds across space. Importantly, we tested musicians with extensive formal training and naïve listeners. Several experiments have reported neurocognitive differences across musically-trained and musically-naïve listeners (Tervaniemi, 2009). Here, we are interested in comparing the auditory processing of musicians and non-musicians during the induction of a beat with distinct spatial features. Even more, some studies suggest that musical training might also modulate how the allocation of attention influences sound localization (e.g. Vuust, Brattico, Seppänen, Näätänen, & Tervaniemi, 2012). Our second experiment thus included the manipulation of attention to explore these possible differences in spatial meter induction.

2. Experiment 1

The first study explores the effect of formal training in music upon the spatial localization of sounds. A good example of the spatial localization of sound in Western music is the disposition of the orchestra players during a concert, where the audience perceives faster and higher-frequency instruments (violins, flute, clarinets…) slightly placed at the left and the metrically-slower and lower-frequency instruments (bass, cello, tuba…) at the right. The beat is normally played by the lowest pitched instruments (Trainor, 2015). It is therefore likely that formal musical training enhances attention to the spatial cues of rhythm, especially to those signaling the beat and the meter. To explore this possibility, we compare musicians (people with formal training in music) and non-musicians (people who never studied music formally nor played any instrument). Several studies have reported structural and functional differences in the brains of musicians (see Schlaug, 2015; Tervaniemi, 2009), but it is still debated up to what point these are a consequence or a prerequisite of music learning. Importantly for the present study, it has been observed that the processing of sound localization varies across listeners depending on their experiences; and musical training seems to contribute to this processing improving binaural hearing, even when attention is driven away (Parbery-Clark, Strait, Hittner, & Kraus, 2013). It is thus possible that these differences may modulate how the brain processes the rhythm of music and could boost the neural processing of beat.

2.1. Methods

2.1.1. Participants

Twenty-four healthy participants (14 females, mean age: 23.37 ± 3.87, age range: 18–33) participated in the present study. No participant was excluded from the analyses. Twelve participants (9 females, mean age: 25.08 ± 3.92) had formal training in music (hereafter: *musicians*): five of them obtained or were pursuing the Superior Degree in Music (comprising at least 14 years of training), and seven of them obtained or were pursuing the Professional Degree in Music (comprising at least 10 years of training). The former degree (known as

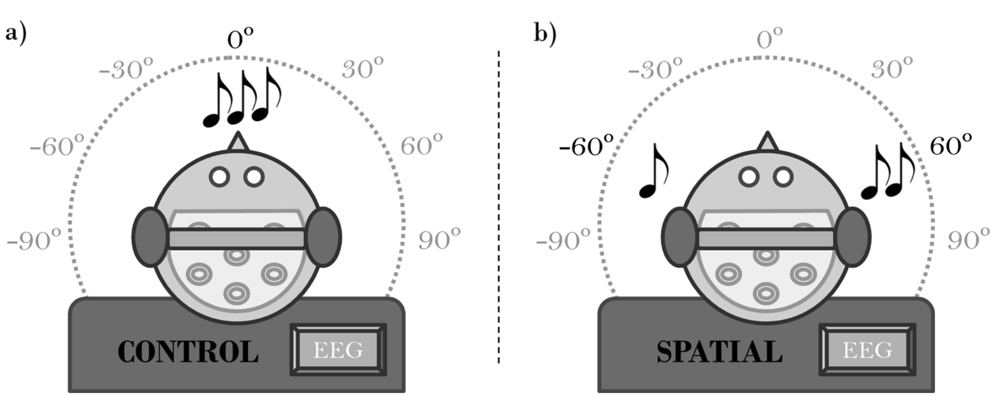

Fig. 1. Schematic representation of participant listening to the experimental conditions. In the control condition (a), the isochronous beat was always presented at 0° (in front of the participant). In the spatial conditions, such as the Spatial 60° (b), the isochronous beat alternated at symmetrical angular positions: the first sound was presented at one side and the two following sounds at the contralateral side, thus following a ternary meter pattern defined over spatial cues.

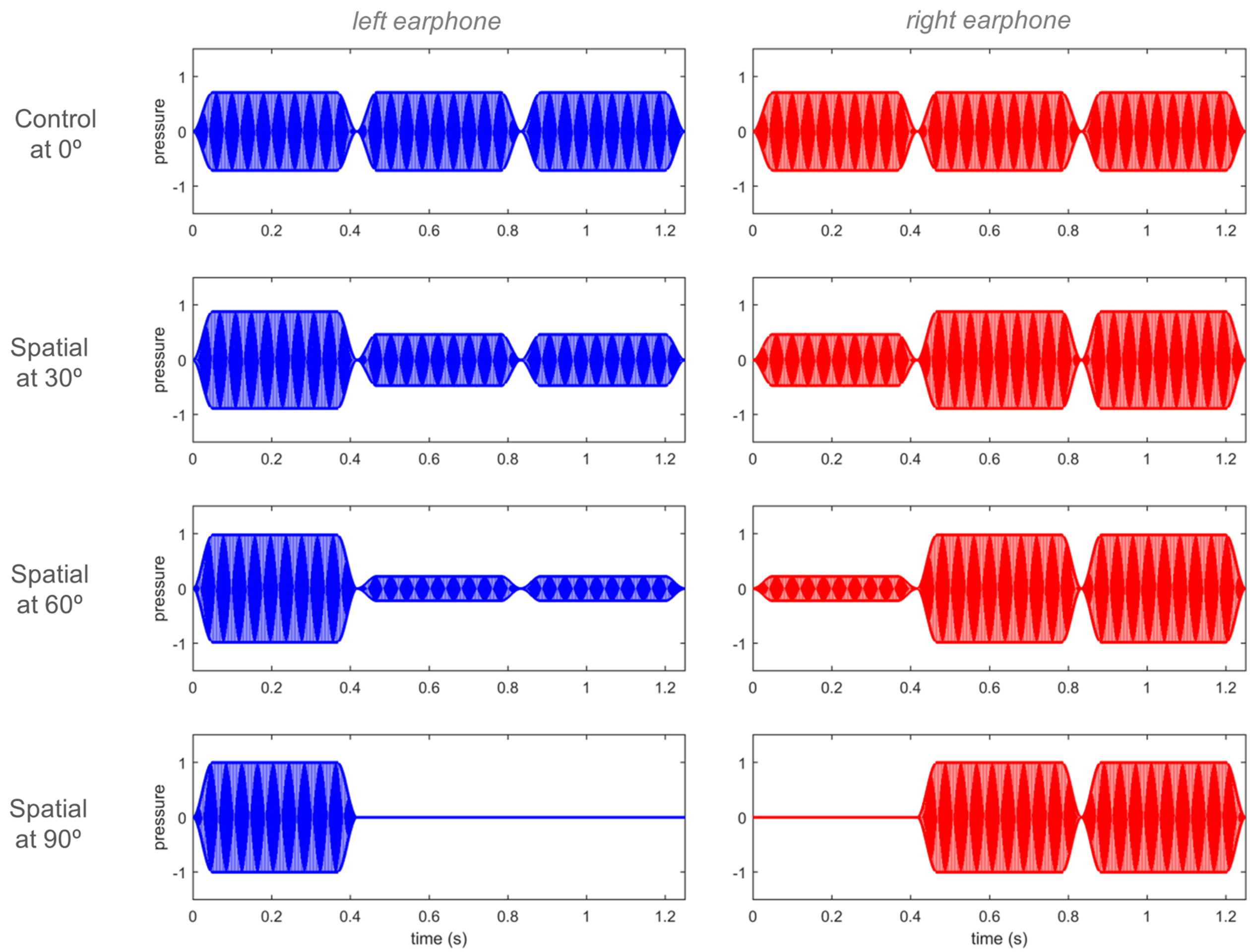

Fig. 2. Modulation of the interaural level and phase differences of three consecutive sounds to emulate distinct angular positions. The columns depict the soundwaves presented at the left and right earphones. Each row shows an instance of the same pitch at the four angular positions: 0°, 30°, 60° and 90°. In the present example, the sound starts with a sound shifted towards the left followed by two sounds shifted towards the right. This pattern was repeated twenty-four times during each 30-second sequence.

"Títol Superior de Música") consists of four years of training equivalent to a Bachelor degree. This can only be accessed at the age of 18 years old after pursuing the Professional Degree (known as "Grau Professional de Música"), which is obtained after at least four years of basic musical training, started at the age of 7–8 years old, and six years of advanced musical training. All the musicians had participated in chamber music ensembles and/or orchestral rehearsals. Three of them also practiced dance for less than five years. The remaining twelve participants (5 females, mean age: 21.67 ± 3.08) did not have any formal training in neither music nor dance (hereafter *non-musicians*) beyond the courses comprised in the obligatory educational system. No participant reported any history of hearing, visual, motor, or psychiatric disorders, and all participants had normal or corrected-to-normal vision. A written informed consent was obtained from all the subjects before the start of the study. All the participants received payment for taking part in the study.

2.1.2. Stimuli

The study comprised four conditions. In each condition, eight sequences of 72 consecutive tones were presented at a frequency of 2.4 Hz (IOI = 416.66 ms). Every sequence lasted 30 s, and each individual tone lasted 416.66 ms, fading in and fading out for 50 ms to avoid the click-artifacts caused by the reproduction of abrupt amplitude changes. To avoid monotony and engage the participants' attention, the fundamental frequency of each 30-second sequence changed. The 72 sounds of a full sequence could all be either at 311, 370, 440 or 523 Hz. These frequencies were chosen because they fall in the comfortable range of human hearing and in the middle of an 88-key piano keyboard. There was a two-second pause after every sequence. All the auditory stimuli were created using Matlab (v.2013, The MathWorks).

The study included a control condition with a constant beat and three spatial conditions in which the sounds were approximately localized at 30°, 60° and 90°, respectively, on the listener's horizontal plane (see Fig. 1). Participants were always presented first with the control condition. In this condition, an isochronous sound at 2.4 Hz was presented at 0°, that is, as if the sound source was placed in front of the participant. The control condition provided a baseline for the neural entrainment to the frequency of the beat. The other three conditions were designed to test whether spatial auditory cues could reliably elicit a ternary meter. The order of presentation of these three spatial conditions followed an increasing angle on the azimuth plane: at 30°, at 60°, and finally at 90°. After the spatial condition at 30° and the spatial condition at 60°, two auditory distractors were presented to the participants. The distractors were the first and second 30-second excerpts of Lux Aeterna by György Ligeti. These distractors did not have any underlying beat and were presented to prevent carry on effects across conditions. With their blurred micropolyphonic rhythmicity, we aimed to counter the neural entrainment to the frequencies of the beat (2.4 Hz) and meter (0.8 Hz).

The ternary meter in the spatial conditions was induced by alternating the spatial location of the sounds, following a pattern in which one sound was presented from one side (e.g. right) and the two following sounds were presented from the other side (e.g. left). Every condition consisted of eight 30-second sequences, and each sequence had twenty-four groups of three alternated sounds. Fig. 2 depicts these three alternated sounds for each condition. The lateralized onset was counterbalanced across sequences: four 30-second sequences starting with a sound at the right side, and four 30-second sequences starting with a sound at the left side. This lateralized alternation occurred at a frequency of 0.8 Hz. Previous studies (Celma-Miralles et al., 2016; Nozaradan, Peretz, Missal, & Mouraux, 2011) showed that the frequency of the beat (2.4 Hz) and its ternary subharmonic (0.8 Hz) fall in the ecological range to elicit neural entrainment in humans.

The progressive modification of the time and the pressure level that the sound reached each ear allowed us to approximately place the sounds at seven spatial degrees: 0°, ± 30°, ± 60° and ± 90°. Based on the comparison of the signals received at the left and right ears, the horizontal position of a source can be computed, for instance, in front of the listener (0°), completely lateralized to the left (−90°) or completely lateralized to the right (90°). The Duplex Theory postulated by Lord Rayleigh (1907) states that, for pure tones, the Interaural Time Differences (ITD) are the most salient cues for the spatial localization of sounds at low frequencies (below 500 Hz), while the Interaural Level Differences (ILD) are mostly used for sounds at higher frequencies. Since we created single-cycle pure tones at frequencies between 311 and 523 Hz, the ITD cues applied to our sounds will be more relevant than the ILD cues. In order to prevent the stimuli from accumulating time delays during the ternary alternations, we converted the ITD values into Interaural Phase Differences, therefore preserving a synchronized sound onset for both channels. In other words, we shifted the phase of the sound sinewave in one channel, and we therefore avoided introducing any time delay between channels. We then computed the exact phase shift for each frequency assuming an approximate time delay of 0.6 ms at 90°, 0.43 ms at 60° and 0.22 ms at 30°. In our model, the ITD values taken from the estimations in Feddersen, Sandel, Teas, and Jeffress (1957) were barely decreased around 10% to compensate for the subsequent application of an amplitude panning.

The use of ILD cues works as an amplitude panning that increases the signal in one channel while it proportionally reduces the signal in the other. This effect emulates and magnifies the distinct spatial localization of the sounds. Our amplitude panning was adapted from Blumlein's optimum configuration designed for a 2.0 stereo system placed at 30° in each side in front of the participant. Since we are presenting the stimuli through headphones, the lateralization of the sounds at 30° is now perceived at 90° by the listener. We applied the tangent law (Bennett, Barker, & Edeko, 1985) to calculate the gains at each channel solving the formula below (Garcia-Vernet, 2017).

$$(1)\ \tan\theta = \frac{gL - gR}{gL + gR}\tan\theta l \rightarrow gL = \sqrt{\frac{1}{1+\left(\frac{-(a-1)}{a+1}\right)^2}} \quad \& \quad gR = \sqrt{1-gL^2} \quad \& \quad a = \frac{\tan\theta}{\tan\theta l}$$

Within the formula, θ stands for the desired spatially-positioned angle, θl stands for the listener position, gL stands for the gain at the left channel, and gR stands for the gain at the right channel. Although this artificially computed model for ILDs does not directly take into account the frequency of the sounds, the obtained amplitude differences between headphones can be interpreted as ILD estimations (Borgo, Soranzo, & Grassi, 2012). The effect of adding ILD on the two differently phase-shifted sound channels might also soften the intracranial lateralization effect of the sounds presented between the two ears (Gelfand, 1982). The combination of ILD and ITDs results in an approximation to the Head-Related Transfer Functions that is applicable to all the participants.

2.1.3. Design and procedure

The study was run in a soundproof room. Participants were seated in a comfortable armchair during the EEG recordings, which lasted an hour (including the preparation of the EEG electrode cap). Participants were instructed to sit comfortably to avoid any body movement, as well as to look at a fixed point in front of them to reduce eye artifacts. The task instructions were simple: "please, listen to the sounds". During the presentation of the stimuli, the experimenter was always out of the room monitoring the experiment. There was a pause after the control condition. During the pause, the experimenter entered in the room and reminded the participant to pay attention to the sounds. This pause was also used to check whether the participant had any problem in hearing the sounds through the headphones.

The auditory stimuli were presented using high-accuracy headphones (Sennheiser HD 435). The use of headphones allowed us to present the same stimuli to all the participants, without the need to take into account inter-individual differences regarding head size, pinnae shape or body position across participants. The sounds were presented at a hearing level of 40 dB, using the matlab Psychophysics Toolbox extensions (Brainard, 1997). After the presentation of all the auditory stimuli, the participant had to answer a brief questionnaire including two questions: (a) did you detect that the sound alternated from right to left? (b) did you detect any pattern of alternation?

All the experimental procedures were approved by the ethical committee from the European Research Council and the Universitat Pompeu Fabra (reference number: 2012/4852/I), and they were carried out in accordance with Spanish and European guidelines.

2.1.4. Electrophysiological recording

The EEG signal was recorded using a BrainAmp amplifier and the BrainVision Analyzer Software package (v.2.0; Brain Products) using an actiCAP with 32 electrodes placed on the scalp (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, C3, Cz, C4, T7, T8, TP9, TP10, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, Oz). Vertical and horizontal eye movements were monitored using two electrodes placed on the right eye: at the outer canthus and at the infra-orbital ridge. Two additional electrodes were placed on the left and right mastoid. The signals were referenced to the FCz online channel and all electrode impedances were kept below 20 kΩ. The signals were amplified and digitized at a sampling rate of 1000 Hz.

2.1.5. Data processing

Preprocessing of the continuous EEG recordings was implemented using BrainVision Analyzer 2.1 (Brain Products GmbH). First, the EEG data was down-sampled to 500 Hz, which does not affect the slow frequencies analyzed in the study. Subsequently, any channel that appeared flat or noisy was interpolated from the surrounding channels via spherical spline interpolation. All the channels were re-referenced to a common average-reference, and the old reference channel FCz was therefore available for the analyses. The signal was filtered using a zero-phase Butterworth filter to remove slow drifts in the recordings with a notch filter (50 Hz), a high pass filter (0.1 Hz, 24 dB/octave) and a low pass filter (30 Hz, 24 dB/octave). Epochs with EEG exceeding either ± 75 µV at any channel, activity < 0.5 µV, or voltage step/sampling > 100 µV within intervals of 200 ms, were automatically detected offline. The eye blinks and muscular movements were removed from the signal by using the Ocular Correction ICA. This correction was based on approximately 300 s of EEG data free of bad intervals. Finally, the filtered EEG data was segmented into 30-second sequences corresponding to the trials of each condition. All further analyses were performed in Matlab and SPSS (version 19, IBM).

For each condition, eight epochs lasting 27.5 s were obtained by removing the first 2.5 s of each sequence. This removal discards the evoked potential related to the stimuli onset and relies on the finding that the steady-state requires several repetitions or cycles to be elicited (Nozaradan et al., 2011). In order to enhance the signal-to-noise ratio and attenuate activities that are not phase locked to the auditory stimuli, the EEG epochs for each participant and condition were averaged across trials. To get the signal's amplitude (in μV), we applied a fast Fourier transform (FFT). Both frequency spectra ranged from 0 to 500 Hz with a frequency resolution of 0.036 Hz. The obtained signal is assumed to correspond to the EEG activity induced by the physical stimuli and the induced meter. However, it may also include residual background noise due to spontaneous activity. Subsequently, we applied the signal-to-noise method used in Nozaradan et al. (2011) to the amplitudes obtained with the FFT. At each frequency bin between 0 and 500 Hz, we subtracted the averaged amplitude of the two surrounding non-adjacent frequency bins, ranging from −0.182 to −0.109 Hz and from 0.109 to 0.182 Hz. The assumption behind this procedure is that, if there is no neural entrainment to specific frequencies, each frequency bin should vary in a similar random way and no peaks should recurrently appear at the same frequency in distinct trials and conditions. This signal-to-noise procedure was applied before the averaging of all the scalp electrodes. Moreover, the fine frequency resolution we obtained allowed us to observe the induced activity centered very concisely in every frequency bin.

Since our approach does not deal with topological effects, we averaged each participant electrodes across the scalp. Clear peaks appeared in the frequency spectra of the four conditions, centered at 0.8, 1.6, 2.4, 3.2, 4 and 4.8 Hz (see Fig. 3). The means of musicians and non-musicians were calculated separately to attest for group differences. For statistical analyses, the amplitudes at each target frequency were submitted firstly to a one-sample *t*-test and secondly to a three-way Mixed Design ANOVA with the within-factors Frequency (0.8, 1.6, 2.4, 3.3, 4, 4.8 Hz) and Condition (Control, Spatial 30°, Spatial 60°, Spatial 90°), and the between-factor Participants (Musicians, non-Musicians). When one of the ANOVA factors was significant, *post hoc* two-tailed *t*-tests with the Bonferroni correction were applied. The significance level was always set at $p < 0.05$.

Taking advantage of the alternating nature of our stimuli, we explored whether there is any hemispheric lateralization in the processing of beat and meter. The "two-component" model by Lerdahl and Jackendoff (1981) proposes that rhythm is processed towards the left hemisphere while the meter towards the right hemisphere. However, there is conflicting evidence supporting this idea: a MEG study recorded activity in the right temporal lobe for subjectively induced binary and ternary meters (Fujioka et al., 2010), an EEG study recorded bilateral frontotemporal cortical activity for meter and rhythms (Kuck, Grossbach, Bangert, & Altenmüller, 2003), and an fMRI study showed activations in left premotor and parietal areas for metrical rhythms (Sakai et al., 1999; see also Platel et al., 1997; Grahn & Brett, 2007; Hong, 2015). In our study, we compare the electrodes from both sides of the scalp to elucidate whether any hemispheric specialization is present for beat and meter. At the same time, this comparison may also shed light onto the idea of a right hemispheric bias for spatial location (Brunetti et al., 2005; Johnson & Hautus, 2010; Kaiser, Lutzenberger, Preissl, Ackermann, & Birbaumer, 2000; Zatorre & Penhune, 2001), and any dominant contralateral processing of spatially located sounds (Fujiki, Riederer, Jousmäki, Mäkelä, & Hari, 2002; Ungan, Yagcioglu, & Goksoy, 2001).

In order to inquire into the lateralization of the auditory mechanisms processing the rhythm of the alternating right and left stimuli, we pursued a second analysis separating the neural signal depending on the stimuli onset location. We averaged separately the four EEG epochs with the stimuli onset on the right and the four EEG epochs with the onset on the left. Subsequently, the same segmentation, FFT and signal-to-noise procedure were applied to both: the right-onset recordings and the left-onset recordings. In order to divide the scalp in two lateralized regions of interest, we separately averaged the right hemisphere electrodes (Fp2, F4, F8, FC2, FC6, C4, T8, TP10, CP2, CP6, P4, P8) and the left hemisphere electrodes (Fp1, F7, F3, FC5, FC1, C3, T7, TP9, CP5, CP1, P7, P3). The central electrodes (Fz, FCz, Cz, Pz, Oz), that are placed in the midline, were removed from the average. For this second statistical analysis, the amplitudes at each target frequency were submitted to a five-way Mixed Design ANOVA with the within-factors Stimuli Onset (Right, Left), Scalp Lateral (Right, Left), Frequency (0.8, 1.6, 2.4, 3.3, 4, 4.8 Hz) and Condition (Control, Spatial 30°, Spatial 60°, Spatial 90°), and the between-factor Participants (Musicians, non-Musicians). When one of the ANOVA factors was significant, *post hoc t*-tests, two-tailed, with the Bonferroni correction were applied. The significance level was always set at $p < 0.05$.

2.2. Results

2.2.1. Frequency analyses of the steady-state evoked potentials

In this study there were four conditions: the control, in which a repetitive sound was placed in front of the participant (at 0°), and three spatial conditions, in which one sound was presented lateralized towards one side and was followed by two consecutive sounds presented in the contralateral side (at 30°, 60° and 90°). After averaging all the scalp electrodes, the EEG signal was converted from the time dimension into the frequency dimension. For each condition, the frequency spectra of each participant, as well as the mean of each group (musicians and non-musicians) are depicted in Fig. 3. The control condition clearly elicited two peaks centered at the frequency of the beat (2.4 Hz) and its harmonic (4.8 Hz). The three spatial conditions also elicited several smaller peaks centered at the frequency of the meter (0.8 Hz) and most of its harmonics (1.6, 3.2 and 4 Hz). There are differences among the amplitudes of all relevant frequencies: the peaks found for the beat and its harmonic are greater than those found for the ternary meter and its harmonics. As Fig. 3 shows, musicians have peaks with more amplitude than non-musicians. A peak at 4 Hz (the fifth harmonic of the meter) is only found for musicians, compared to non-musicians. Besides, non-musicians show a smaller peak at 0.8 Hz (the frequency of the meter) for the 30° conditions than for the 60° and 90° conditions. Importantly, the frequency spectrum of the auditory distractor used between the spatial conditions does not reveal any observable peak at our frequencies of interest.

As Nozaradan et al. (2011), we assume that, if no steady-state evoked potentials are present, all the frequencies of the spectra may vary similarly. After the subtraction of the mean of adjacent frequency bins is applied to each frequency bin, all the random peaks should tend towards zero (2011). To discover the presence of beat- and meter-related peaks in our frequency spectra, we took all participants' magnitudes centered at our frequencies of interest. For all four conditions, *t*-tests against zero revealed that beat-related SSEPs were elicited at 2.4 Hz (all $p < .001$) and at 4.8 Hz (all $p < .001$). For the three spatial conditions, t-tests against zero showed that meter-related SSEPs were elicited at 0.8 Hz (all $p < .014$), at 1.6 Hz (all $p < .001$), at 3.2 Hz and (all $p < .001$), and at 4 Hz (all $p < .012$). If the t-tests against zero are separately applied to each group (see Table 1), similar significant peaks appear for the same frequencies and conditions in musicians. In contrast, four of these peaks become non-significant in non-musicians: at 0.8 Hz (for the Spatial 60° condition, $t(11) = 1.466$, $p = .170$) and at 4 Hz (for the Spatial 30°, $t(11) = 1.326$, $p = .212$; for the Spatial 60°, $t(11) = 2.169$, $p = .053$; and the Spatial 90°, $t(11) = 1.027$, $p = .327$). If the Bonferroni correction is applied to the twelve tests comprised in each frequency, the alpha drops to the conservative level of significance $p < .004$. At this level of significance (see the asterisks in Table 1), the non-musicians' peaks for the spatial conditions at 0.8, 1.6, 3.2 and 4 Hz cannot be considered different from zero. Some of the non-musicians' peaks of the Spatial Conditions also cannot pass the correction: at 0.8 Hz (for the Spatial 60° and 90°), at 1.6 Hz (for the Spatial 30°), and at 4 Hz (for the Spatial 90°). Despite this, when the peaks are analyzed all together, the only non-significant peak is found at 0.8 Hz (for the Spatial 60°) and at 4 Hz (for the Spatial 90°).

A three-way mixed design ANOVA was applied to the peak magnitudes with the within-factors Frequency (0.8, 1.6, 2.4, 3.2, 4, 4.8 Hz) and Condition (Control, Spatial 30°, Spatial 60°, Spatial 90°), and the between-factors Group (Musicians, non-Musicians). When the

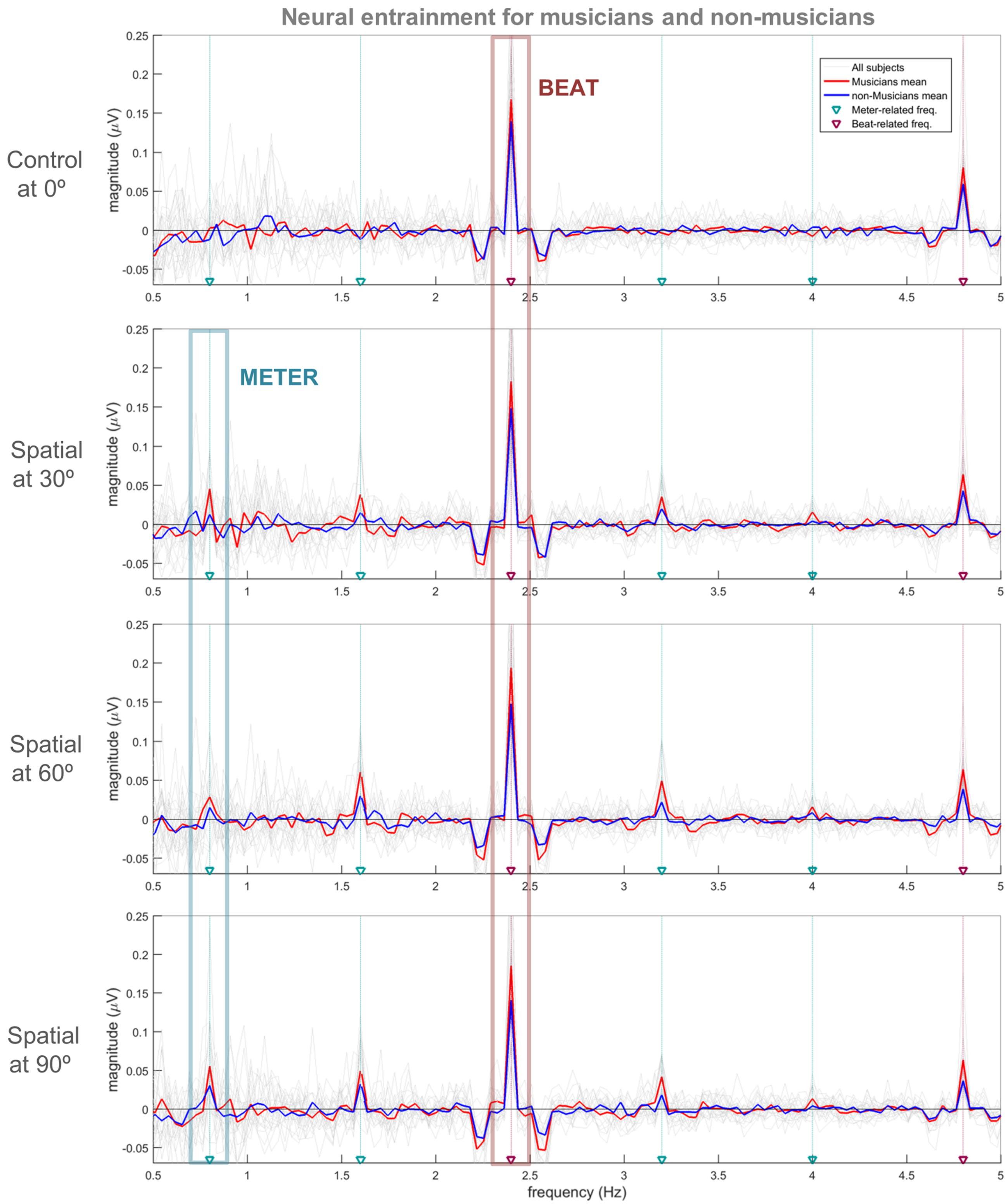

Fig. 3. Amplitude spectra of all the participants across the four conditions. The average of the musicians' magnitudes is depicted in red and the average of the non-musicians' magnitudes is depicted in blue. The y-axis represents the amplitudes in microvolts, and the x-axis represents the frequency continuum. Green triangles stand for the frequencies related to the ternary meter (0.8 Hz) and its harmonics (1.6, 3.2 and 4 Hz). Red triangles stand for the frequencies related to the beat (2.4 Hz) and its harmonic (4.8 Hz). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

assumption for sphericity was violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. The analyses of variance showed a main effect of Group, $F_{(1,22)} = 4.334$, $p = .049$, $\eta^2 = 0.165$, a main effect of Frequency, $F_{(1.512,33.262)} = 71.545$, $p < .001$, $\eta^2 = 0.765$, a main effect of Condition, $F_{(3,66)} = 18.368$, $p < .001$, $\eta^2 = 0.455$, and an interaction between Frequency and Condition, $F_{(7.336,161.400)} = 6.003$, $p < .001$, $\eta^2 = 0.214$. The interaction between Group and Condition showed a tendency towards significance, $F_{(3,66)} = 2.666$, $p < .055$, $\eta^2 = 0.108$.

Post hoc pairwise comparisons with the Bonferroni alpha correction

Table 1
One-sample *t*-tests against zero for all participants, musicians and non-musicians at each frequency of interest and across conditions.

		df	Control 0°		Spatial 30°		Spatial 60°		Spatial 90°	
			t	*p*	*t*	*p*	*t*	*p*	*t*	*p*
0.8 Hz	All	23	−0.582	.566	**4.557***	< .001	**2.661**	0.014	**3.539***	.002
	M	11	0.283	.783	**4.529***	< .001	**2.209**	0.049	**2.646**	.023
	nM	11	−1.208	.252	**2.732**	.020	1.466	0.171	**2.505**	.029
1.6 Hz	All	23	**−4.578***	< .001	**3.818***	< .001	**6.007***	< 0.001	**4.815***	< .001
	M	11	**−3.055**	.011	**3.016**	.012	**5.608***	< 0.001	**5.158***	< .001
	nM	11	**−3.311**	.007	**3.532**	.005	**3.377**	0.006	**2.306**	.042
2.4 Hz	All	23	**9.857***	< .001	**9.360***	< .001	**9.450***	< 0.001	**8.772***	< .001
	M	11	**6.340***	< .001	**5.985***	< .001	**5.854***	< 0.001	**5.275***	< .001
	nM	11	**8.335***	< .001	**8.277***	< .001	**11.210***	< 0.001	**12.779***	< .001
3.2 Hz	All	23	−0.334	.741	**5.311***	< .001	**5.738***	< 0.001	**5.886***	< .001
	M	11	−1.017	.331	**4.636***	< .001	**4.742***	< 0.001	**5.046***	< .001
	nM	11	0.280	.784	**2.989**	.012	**4.815***	< 0.001	**4.543***	< .001
4 Hz	All	23	−0.670	.510	**3.656***	.001	**4.130***	< 0.001	**2.742**	.012
	M	11	**−2.510**	.029	**4.054***	.002	**3.682***	0.004	**2.798**	.017
	nM	11	1.078	.304	1.326	.212	2.169	0.053	1.027	.327
4.8 Hz	All	23	**7.481***	< .001	**7.158***	< .001	**8.023***	< 0.001	**7.319***	< .001
	M	11	**4.537***	< .001	**5.090***	< .001	**6.234***	< 0.001	**5.224***	< .001
	nM	11	**10.641***	< .001	**5.795***	< .001	**6.285***	< 0.001	**9.192***	< .001

For each condition ('Control', 'Spatial 30°', 'Spatial 60°', 'Spatial90°') and frequency of interest (0.8, 1.6, 2.4, 3.2, 4, 4.8 Hz), the degrees of freedom (df), the *t*-statistic and its *p*-value is reported. The results for all participants ('All') appear in gray background, while the results for musicians ('M') and non-musicians ('nM') appear in white background. The *t*-statistics below a significance level of 0.05 appear in bold. The *t*-statistics that pass the Bonferroni correction considering the 12 tests of each frequency are marked with an asterisk (significance level: $p < .004$). For the Spatial conditions, the *t*-statistics above significance level appear underlined, all cases for non-musicians. Note that for the Control condition, the *t*-statistics at meter-related frequencies are either significantly negative or non-significant.

Table 2
Post hoc pairwise comparisons from the three-way mixed design ANOVA with the Bonferroni adjustment for multiple comparisons.

	Condition	Condition	Mean difference	Standard Error	P-values
at 0.8 Hz	Spt. 30°	Cnt. 0°	0.034	0.008	.003**
	Spt. 60°		0.026	0.009	.038*
	Spt. 90°		0.047	0.013	.007**
at 1.6 Hz	Spt. 30°	Cnt. 0°	0.037	0.007	< .001***
	Spt. 60°		0.056	0.007	< .001***
	Spt. 90°		0.052	0.009	< .001***
at 2.4 Hz	Spt. 30°	Cnt. 0°	0.012	0.010	1
	Spt. 60°		0.017	0.008	.270
	Spt. 90°		0.009	0.010	1
at 3.2 Hz	Spt. 30°	Cnt. 0°	0.029	0.007	.003**
	Spt. 60°		0.037	0.007	< .001***
	Spt. 90°		0.031	0.007	.001**
at 4.0 Hz	Spt. 30°	Cnt. 0°	0.012	0.003	.003**
	Spt. 60°		0.014	0.004	.011*
	Spt. 90°		0.011	0.004	.007**
at 4.8 Hz	Spt. 30°	Cnt. 0°	−0.016	0.006	.104
	Spt. 60°		−0.019	0.007	.104
	Spt. 90°		−0.020	0.007	.041*

For the six frequencies of interest, the mean differences between each spatial condition and the control, their standard error and corrected *p*-values are reported. Significance level is always kept below 0.05. The beat-related frequencies (2.4, 4.8 Hz) appear in gray background, while the meter-related frequencies (0.8, 1.6, 3.2, 4 Hz) appear in white background. The label ' Cnt. 0° ' stands for the Control condition with the beat at 0°, the labels ' Spt. 30° ', ' Spt. 60° ' and ' Spt. 90° ', stand for the Spatial conditions with the alternated ternary meter at 30°, 60° and 90°, respectively. Note that the only significant comparison at a beat-related frequency is at 4.8 Hz, where the Spatial 90° condition show smaller amplitudes than the Control condition. *$p < .05$, **$p < .01$, ***$p < .001$

revealed that (i) the amplitudes of the musicians were greater than those of non-musicians (MD = 0.02, $p = .049$), and that (ii) the amplitudes of the three spatial conditions were higher than the amplitudes of the control condition at meter-related frequencies (0.8, 1.6, 3.2, 4 Hz), but not at beat-related frequencies (2.4, 4.8 Hz). The only exception was the amplitudes at 4.8 Hz of the Spatial 90° condition, which were significantly smaller than the amplitudes of the control (MD = 0.02, $p = .041$). The results of each pairwise comparison are reported in Table 2 and depicted in Fig. 4. Across conditions, the highest amplitudes always appeared at the beat frequency 2.4 Hz ($p < .001$), which are directly related to the processing of the external auditory stimuli.

The pairwise comparisons of the non-significant interaction between Group and Condition suggest that the amplitudes of musicians were significantly greater than those of non-musicians in the Spatial 30° condition (MD = 0.023, $p = .018$) and the Spatial 60° condition (MD = 0.025, $p = .036$), but not in the Spatial 90° condition (MD = 0.024, $p = .065$) nor in the Control condition (MD = 0.008, $p = .370$). Such pattern of results suggests that the differences between groups tend to disappear when the ternary meter is more evident, that is, when the sounds alternate at 90° in the azimuth plane (see Fig. 5). If the three Spatial conditions are compared to the Control condition within each group, musicians show greater amplitudes in the Spatial 30° condition (MD = 0.025, $p < .001$), the Spatial 60° condition (MD = 0.030, $p < .001$) and the Spatial 90° condition (MD = 0.030, $p < .001$). Non-musicians only show greater amplitudes in the Spatial 60° condition (MD = 0.014, $p = .042$), but not in the Spatial 30° condition (MD = 0.010, $p = .277$) nor in the Spatial 90° condition (MD = 0.014, $p = .134$). Thus, non-musicians might be less familiar and accurate in entraining to the frequency of the ternary meter (despite clearly succeeding in the Spatial 60° condition). In contrast, musicians might have an easier access to the metrical structure and therefore a more accurate entrainment to the ternary meter elicited by the auditory spatial cues at all the angular positions.

For the analysis regarding the stimuli onset location and the lateralization of the electrodes, a five-way mixed design ANOVA was applied. The analyses of variance showed a main effect of Group, $F_{(1,22)} = 4.324$, $p = .049$, $\eta^2 = 0.164$, a main effect of Frequency, $F_{(1.645,36.186)} = 61.539$, $p < .001$, $\eta^2 = 0.737$, a main effect of Condition, $F_{(3,66)} = 20.545$, $p < .001$, $\eta^2 = 0.483$, and an interaction between Frequency and Condition, $F_{(7.205,158.520)} = 5.901$, $p < .001$, $\eta^2 = 0.211$. Since these results do not differ from those in the three-way mixed design ANOVA, we avoid reporting here the pairwise comparisons. Our intention was to explore effects or interactions between the location of the downbeat and any lateralized processing in the scalp. However, no main effects were found regarding the stimuli onset

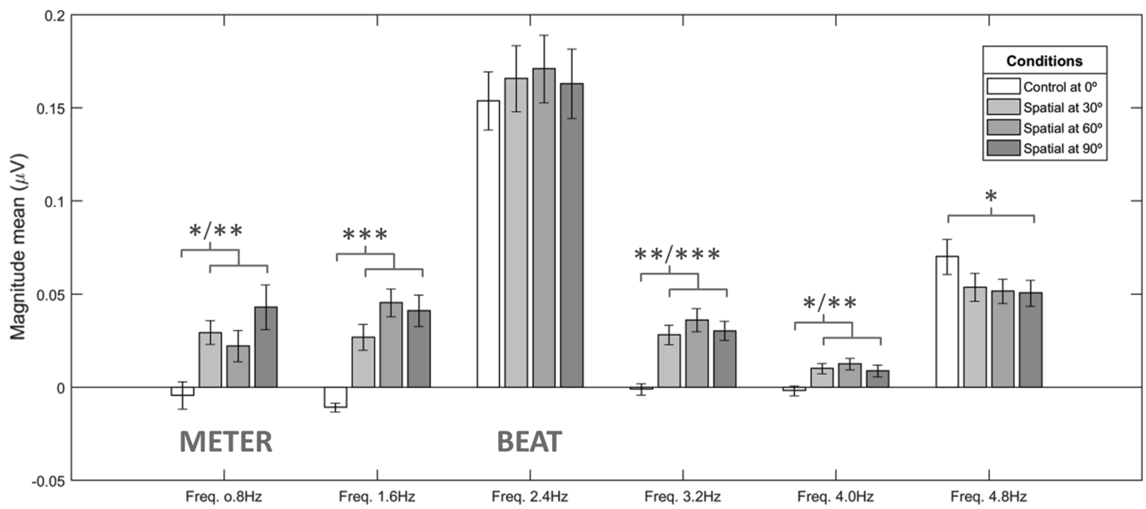

Fig. 4. Mean of the amplitudes at meter-related (0.8, 1.6, 3.2 and 4 Hz) and beat-related (2.4 and 4.8 Hz) target frequencies collapsed across all participants. Standard error bars are depicted. At all the meter-related frequencies, the magnitudes of the Control condition are significantly lower than the magnitudes of the Spatial conditions, regardless of the angle. The significance level is marked with asterisks: * indicates $p < .05$, ** indicates $p < .01$ and *** indicates $p < .001$.

location or the lateralization of the electrodes (please see Supplementary Material for topographical plots displaying the scalp distribution of the responses). This lack of effects on lateralization indicates that our recordings of the neural entrainment show (i) no differences between the stimuli starting on the right or on the left, (ii) no differences between the hemispheres during the processing of the metrical structure, and (iii) no differences between the ipsilateral and contralateral processing of the sounds, likely due to the right-left alternating nature of the stimuli.

2.2.2. Questionnaires

At the end of the study, both groups of participants answered a few questions in an online formulary. We asked them whether they noticed that the sound was coming from different locations. All the participants reported to have detected these differences regarding the sound source. This confirmed that they all detected the auditory spatial alternation of the sounds caused by the modification of the binaural cues of the sound. In another question, we wanted to know whether participants were able to identify what kind of rhythmic pattern was following the lateralized alternation of the sounds. Ten out of the twelve musicians explicitly reported that the pattern followed by the sounds was the ternary meter (waltz-like, ¾ measure), while only three of the twelve non-musicians mentioned this ternary alternation. In addition, two musicians reported to have perceived the downbeat of the ternary meter as longer in duration or lower in pitch. This is interesting because these acoustic modulations often mark the downbeat of a waltz in actual music (London, 2012). We can therefore speculate that the previous experience with downbeat stereotypes (i.e. accented longer tones normally marked by lower pitches) could have modulated the perception of the spatially-located sounds. Non-musicians gave more diverse answers regarding the pattern of the alternation. They mentioned that the pattern consisted of a one-to-one alternation, a three-to-one alternation, or even that the sound was alternating from ear to ear at every second. We may keep in mind that musicians had the advantage of having a term to label the ternary meter alternation, which might be unknown to some non-musicians. Moreover, the two musicians who did not explicitly mention the ternary pattern may have noticed it, but they just answered to our open-question saying that they had detected the rhythmic pattern all the time. In future studies, we should consider asking multiple-choice questions to force participants to give a concise answer and ensure that all answers are truly comparable. Although we did not explicitly control for micro-movements during the experiment, none of

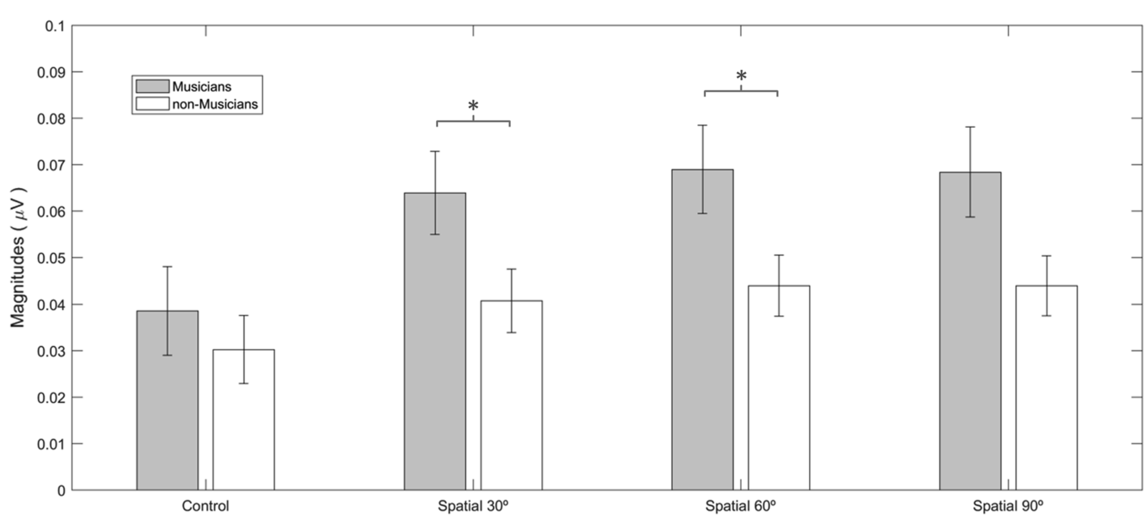

Fig. 5. Neural responses of musicians and non-musicians across conditions. The six frequencies of interest are averaged at each bar and their standard error bars are depicted. The pairwise comparison of the non-significant interaction between Group and Condition revealed that musicians tended to have higher magnitudes in the Spatial 30° and 60° conditions compared to non-musicians. The asterisk indicates the significant level: $p < .05$.

the participants reported to have been tapping while listening to the sounds.

2.3. Discussion

In this first study, we aimed to answer two main questions: (i) how the brain interacts with spatially separated sounds and (ii) to what extent formal training in music shapes the perception of beat and meter in the spatial domain. The results stemming from the frequency-tagging analyses support the idea that the spatial location of sounds interacts with the organization of rhythmic events and music expertise. First, the peaks we observed at metrically-related frequencies correspond to the grouping of three sounds that repetitively alternate from one side to the other. This provides evidence for neural entrainment to a spatially-defined ternary meter. Second, the significant differences in amplitude between the peaks of musicians and non-musicians indicate that formal training in music could facilitate the induction of rhythmic structures defined over spatial information. The neural entrainment of musicians is enhanced for the frequencies of both beat and meter. Finally, our recordings have not captured any lateralization effect for the stimuli location nor the processing of meter.

3. Experiment 2

We ran a second experiment to control for the role of attention in the induction of meter by spatial cues. In this experiment, the participants were presented with the spatially-located sounds while they were watching muted episodes of The Pink Panther. As in the previous experiment, we tested musicians and non-musicians. This is important because there are differences between the auditory skills of jazz, rock and classical musicians compared to non-musicians even at the pre-attentional level (Vuust et al., 2012). For instance, when attention is directed towards the visual modality, changes in source-location elicit enhanced mismatch negativities in jazz musicians (Vuust et al., 2012) and amateur rock musicians (Tervaniemi, Castaneda, Knoll, & Uther, 2006). This suggests that spatial features are particularly relevant for these performers, perhaps because this helps them to either communicate with each other during live music performances or to deal with the electronic sound amplifiers distributed on the scene. Interestingly, spatial sound information seems to be also crucial for expert classic-music conductors: it has been observed that their neural mechanisms are activated regardless of their attentional focus on the space periphery (Münte, Kohlmetz, Nager, & Altenmüller, 2001; Nager, Kohlmetz, Altenmüller, Rodriguez-Fornells, & Münte, 2003). Thus, in the present experiment, we aimed to disentangle whether the EEG responses elicited by the sounds were due to a low-level automatic processing of the auditory spatial contrasts or whether they depended on a high-level processing of the sounds. In other words, if we do not capture neural entrainment to meter, we can speculate that the organization of sounds into cyclical patterns requires the combination of bottom-up and top down mechanisms.

3.1. Methods

3.1.1. Participants

Thirty-two healthy participants (17 females, mean age: 23.09 ± 4.22, age range: 18–33) participated in the present study. No participant was excluded from the analyses. Twelve participants (9 females, mean age: 25.19 ± 4.09) had formal training in music (hereafter: *musicians*): five of them obtained or were pursuing the Superior Degree in Music, nine of them obtained or were pursuing the Professional Degree in Music, one of them obtained the Cultural Degree in Traditional Music and two of them pursued at least 10 years of musical training in the musical educational system of their country. All the musicians had participated in chamber music ensembles and/or orchestral rehearsals. Eight of them also practiced dance. The remaining sixteen participants (8 females, mean age: 21 ± 3.286) did not have any formal training in neither music nor dance (hereafter *non-musicians*). Participants did not report any history of hearing, visual, motor, or psychiatric disorders. All participants had normal or corrected-to-normal vision. They signed a written informed consent and received payment for taking part in the study.

3.1.2. Stimuli

The auditory stimuli were identical to the ones used in Experiment 1 (see Section 2.1.2). Here, the sounds were presented together with a visual distractor that consisted of twelve consecutive muted episodes of The Pink Panther. The presentation of the distractor movies was accelerated 25% in terms of frames per second.

3.1.3. Design and procedure

As in Experiment 1, participants sat in a soundproof room during the EEG recordings. They were asked to watch a silent video while the auditory stimuli were presented through the headphones. After the EEG recordings, there was a short questionnaire about the auditory sounds and a four-alternative forced-choice test about The Pink Panther episodes, to control that participants were paying attention to the movie.

3.1.4. Electrophysiological recording

The EEG signal was recorded as reported in Experiment 1 (see Section 2.1.4).

3.1.5. Data processing

We implemented the same preprocessing and frequency analyses of the EEG signal as stated in Experiment 1 (see Section 2.1.5). In addition to the previous statistical analyses, we added a comparison between Experiment 1 and Experiment 2 to attest the role of attention in spatial meter induction. To properly contrast across experiments, we subtracted the amplitudes of the Control condition from each Spatial condition. This subtraction allowed us to directly compare the modulation of neural entrainment among the Spatial conditions focusing on attention and musical training.

3.2. Results

3.2.1. Frequency analyses of the steady-state evoked potentials

As in Experiment 1, we took all participants' magnitudes centered at our frequencies of interest and applied *t*-tests against zero. This revealed that beat-related SSEPs were elicited at 2.4 Hz (all $p < .001$) and at 4.8 Hz (all $p < .001$). For the three spatial conditions, t-tests against zero showed that meter-related SSEPs were elicited at 0.8 Hz (all $p < .019$), at 1.6 Hz (all $p < .0.46$), at 3.2 Hz and (all $p < .032$), but only in the Spatial 60° and 90° conditions at 4 Hz (both $p < .05$). We computed *t*-tests against zero separately to each group (see Table 3). Only musicians showed significant peaks at the frequency of the meter. If the Bonferroni correction is applied to the twelve tests comprised in each frequency, the alpha drops to a more conservative level of significance $p < .004$, turning most of the amplitudes into non-significantly different from zero. For instance, at 0.8 Hz, only musicians surpass this level of significance in the Spatial 60° condition (see the asterisks in Table 3).

A three-way mixed design ANOVA was applied to the peak magnitudes with the within-factors Frequency (0.8, 1.6, 2.4, 3.2, 4, 4.8 Hz) and Condition (Control, Spatial 30°, Spatial 60°, Spatial 90°), and the between-factors Group (Musicians, non-Musicians). When the assumption for sphericity was violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. The analyses of variance showed a main effect of Frequency, $F_{(1.914,57.419)} = 84.0.848$, $p < .001$, $\eta^2 = 0.739$, a main effect of Condition, $F_{(3,90)} = 5.470$, $p = .002$, $\eta^2 = 0.154$, and an interaction between Frequency and Condition, $F_{(8.022,240.653)} = 4.324$, $p < .001$, $\eta^2 = 0.126$. There was no main effect of Group, $F_{(1,30)} = 2.372$, $p = .134$, $\eta^2 = 0.073$. *Post hoc*

Table 3
One-sample *t*-tests against zero for all the participants, musicians and non-musicians at each frequency of interest and across conditions.

Without attention		df	Control 0°		Spatial 30°		Spatial 60°		Spatial 90°	
			t	*p*	*t*	*p*	*t*	*p*	*t*	*p*
o.8 Hz	All	31	−0.307	.761	**2.470**	.019	**2.598**	.014	**2.896**	.007
	M	15	**−2.144**	.049	**2.492**	.025	**3.355***	.004	**2.423**	.029
	nM	15	1.067	.303	1.182	.256	0.633	.536	1.627	.125
1.6 Hz	All	31	−1.198	.240	**4.280***	< .001	**3.524***	.001	**2.079**	.046
	M	15	−1.169	.261	**2.570**	.021	1.769	.097	0.665	.561
	nM	15	−0.410	.688	**3.453***	.004	**3.161**	.006	2.122	.051
2.4 Hz	All	31	**9.925***	< .001	**10.089***	< .001	**11.100***	< .001	**11.143***	< .001
	M	15	**8.732***	< .001	**8.328***	< .001	**8.991***	< .001	**8.747***	< .001
	nM	15	**6.709***	< .001	**6.713***	< .001	**7.742***	< .001	**7.166***	< .001
3.2 Hz	All	31	−0.151	.881	**2.251**	.032	**4.488***	< .001	**3.229***	.003
	M	15	−1.055	.308	1.139	.273	**3.324**	.005	**2.660**	.018
	nM	15	0.961	.352	**2.334**	.034	**3.051**	.008	1.840	.086
4 Hz	All	31	−0.960	.345	1.855	.073	**3.025**	.005	**3.135***	.004
	M	15	−0.702	.493	1.808	.091	1.984	.066	1.131	.276
	nM	15	−0.632	.537	0.578	.572	**2.263***	.039	**4.301***	.001
4.8 Hz	All	31	**12.370***	< .001	**9.853***	< .001	**9.194***	< .001	**8.254***	< .001
	M	15	**9.093***	< .001	**5.651***	< .001	**5.490***	< .001	**5.235***	< .001
	nM	15	**8.626***	< .001	**8.781***	< .001	**7.927***	< .001	**7.086***	< .001

For each condition ('Control', 'Spatial 30°', 'Spatial 60°', 'Spatial90°') and frequency of interest (0.8, 1.6, 2.4, 3.2, 4, 4.8 Hz), the degrees of freedom (df), the *t*-statistic and its *p*-value is reported. The results for all participants ('All') appear in gray background, while the results for musicians ('M') and non-musicians ('nM') appear in white background. The *t*-statistics below a significance level of 0.05 appear in bold. The *t*-statistics that pass the Bonferroni correction considering the 12 tests of each frequency are marked with an asterisk (significance level: $p < .004$). Note that for the Control condition, the *t*-statistics at meter-related frequencies are either significantly negative or non-significant.

pairwise comparisons with the Bonferroni alpha correction revealed that, compared to the control, there were higher amplitudes at 0.8 Hz for the Spatial 90° (MD = 0.02, $p = .028$), at 1.6 Hz for the Spatial 30° (MD = 0.02, $p = .001$) and Spatial 60° (MD = 0.03, $p = .016$), at 3.2 Hz for the Spatial 60° (MD = 0.02, $p = .005$), and at 4 Hz for the Spatial 60° (MD = 0.01, $p = .031$) and Spatial 90° (MD = 0.01, $p = .020$). The amplitudes at 4.8 Hz were lower for the Spatial 60° (MD = 0.02, $p = .038$) and Spatial 90° (MD = 0.03, $p > .001$). The amplitudes of the beat-related frequencies were higher than the amplitudes of the meter-related frequencies at all conditions (all $p < .001$) except for the Spatial 90°, in which the amplitudes at 4.8 Hz were not higher than those at 0.8 Hz (MD = 0.03, $p = .087$). Finally, the amplitudes at 2.4 Hz were always higher than those at 4.8 Hz ($p < .001$; see Fig. 6).

Next, we compared the results from Experiment 1 and 2. To better observe the modulation of the peaks in the frequencies of interest, we subtracted the amplitudes of the control to the amplitudes of each spatial condition. A four-way mixed design ANOVA was applied to the subtracted values with the within-factors Frequency (0.8, 1.6, 2.4, 3.2, 4, 4.8 Hz) and Condition (Spatial 30°, Spatial 60°, Spatial 90°), and the between-factors Group (Musicians, non-Musicians) and Attention (Sounds, Video). There was a main effect of Group, $F_{(1,52)} = 10.661$, $p = .002$, $\eta^2 = 0.170$, a main effect of Attention, $F_{(1,52)} = 13.827$, $p < .001$, $\eta^2 = 0.210$, and a main effect of Frequency, $F_{(3.834,199.363)} = 24.331$, $p < .001$, $\eta^2 = 0.319$. The *post hoc* pairwise comparisons with the Bonferroni alpha correction revealed greater increases in the peaks in two cases (see Fig. 7): when participants were musicians (MD = 0.01, $p = .002$) and when participants put attention to the sounds (MD = 0.01, $p > .001$). The comparison between the frequencies of interest revealed greater increases at 0.8 Hz compared to 2.4 Hz (MD = 0.02, $p = .005$), 4 Hz (MD = 0.02, $p = .024$) and 4.8 Hz (MD = 0.05, $p < .001$); at 1.6 Hz compared to 2.4 Hz (MD = 0.03, $p = .002$), 4 Hz (MD = 0.02, $p < .001$) and 4.8 Hz (MD = 0.05, $p < .001$); at 2.4 Hz compared to 4.8 Hz (MD = 0.02, $p = .008$); at 3.2 Hz compared to 4.8 Hz (MD = 0.04, $p < .001$); and at 4 Hz compared to 4.8 Hz (MD = 0.03, $p < .001$).

3.2.2. Questionnaires

At the end of the study, participants were asked about the movie and answered the same two questions of Experiment 1. Surprisingly, two non-musicians reported that they did not notice that the sound was coming from different locations. Four musicians and all the non-musicians did not explicitly identify the rhythmic pattern as a ternary structure. However, most participants reported to have noticed the cyclical alternation of the sounds following a repetitive pattern. The average of correct responses about the movie was 86.47% ± 18.43 for non-musicians and 97. 62% ± 5.48 for musicians. Participants thus focused on the visual distractor.

3.3. Discussion

In this second experiment, we explored whether the low-level features of the stimuli are enough to trigger neural entrainment to the ternary meter defined in the spatial domain. We observed that the peaks at the frequency of the ternary meter were only consistent in the group of musicians. Interestingly, significant differences observed across experiments suggest a positive modulation of neural entrainment due to attention.

Our experiment directed the participants' attention to the visual modality. The muted videos aimed to prevent top-down influences in auditory perception. Silent movies were also used in other studies comparing musicians and non-musicians, which revealed that musical training enhances pre-attentive auditory skills in language and music processing (Parbery-Clark et al., 2013; Tervaniemi et al., 2009) and that spatial features are particularly relevant for jazz musicians (Vuust et al., 2012) and amateur rock musicians (Tervaniemi et al., 2006) compared to non-musicians. Unfortunately, our 3-way mixed design ANOVA did not reveal differences between musicians and non-musicians regarding neither the frequencies of interest nor the spatial conditions. However, the *t*-tests against zero (see Table 3) suggest that musicians showed more consistent peaks at the meter frequency (0.8 Hz) in all spatial conditions, whereas non-musicians actually showed more consistent peaks at the frequencies of its first and fourth harmonics (1.6 and 4 Hz) in the 60° condition. At the third harmonic (3.2 Hz), the peaks were more consistent for non-musicians in the Spatial 30° and for musicians in the Spatial 90°, but similar in the Spatial 60°. It is an open question the reason why the peaks of the musicians and non-musicians varied at the frequencies of the meter-harmonics in this way. One could argue

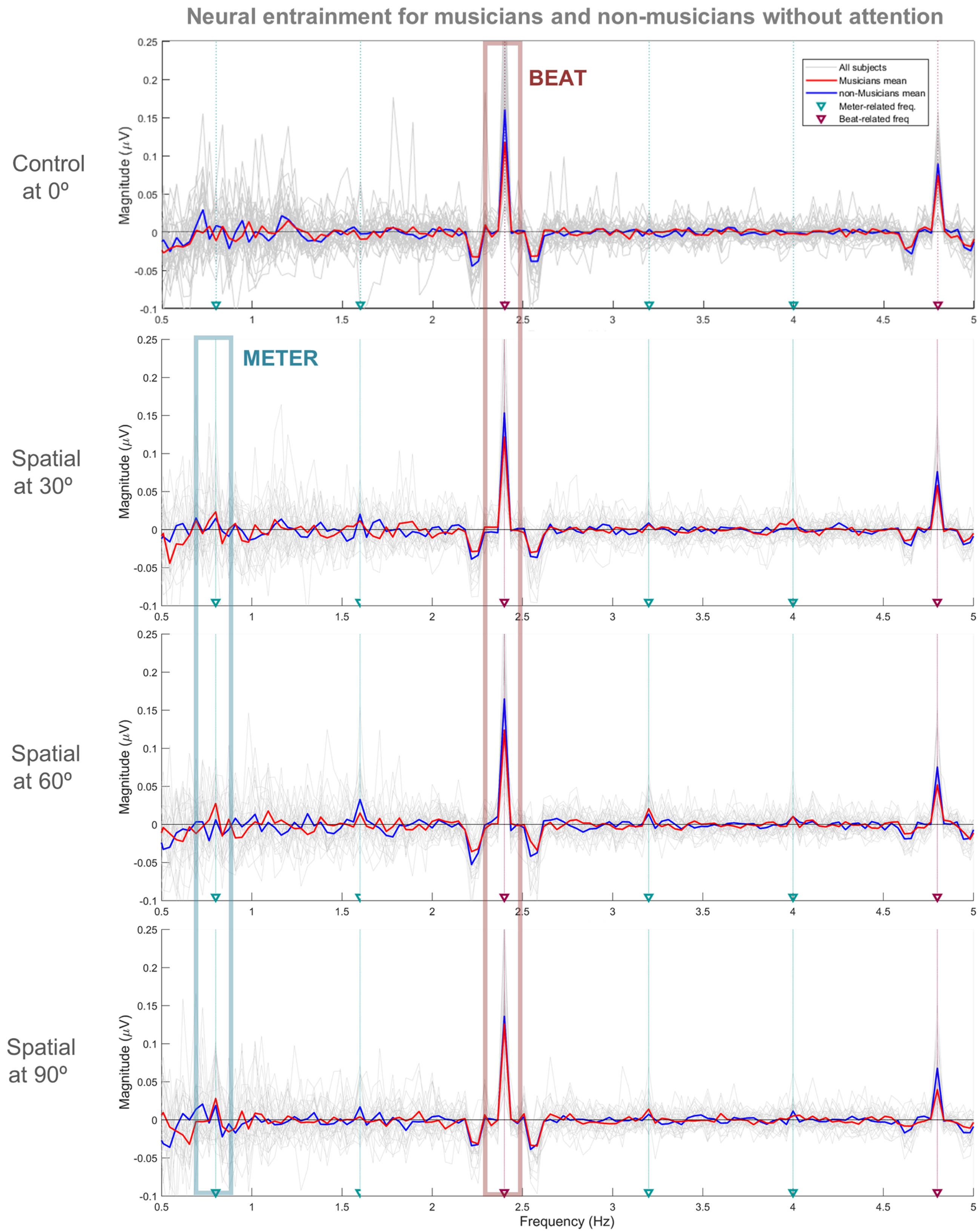

Fig. 6. Amplitude spectra of all the participants across the four conditions while watching a movie. The average of the musicians' magnitudes is depicted in red and the average of the non-musicians' magnitudes is depicted in blue. The y-axis represents the amplitudes in microvolts, and the x-axis represents the frequency continuum. Green triangles stand for the frequencies related to the ternary meter (0.8 Hz) and its harmonics (1.6, 3.2 and 4 Hz). Red triangles stand for the frequencies related to the beat (2.4 Hz) and its harmonic (4.8 Hz). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

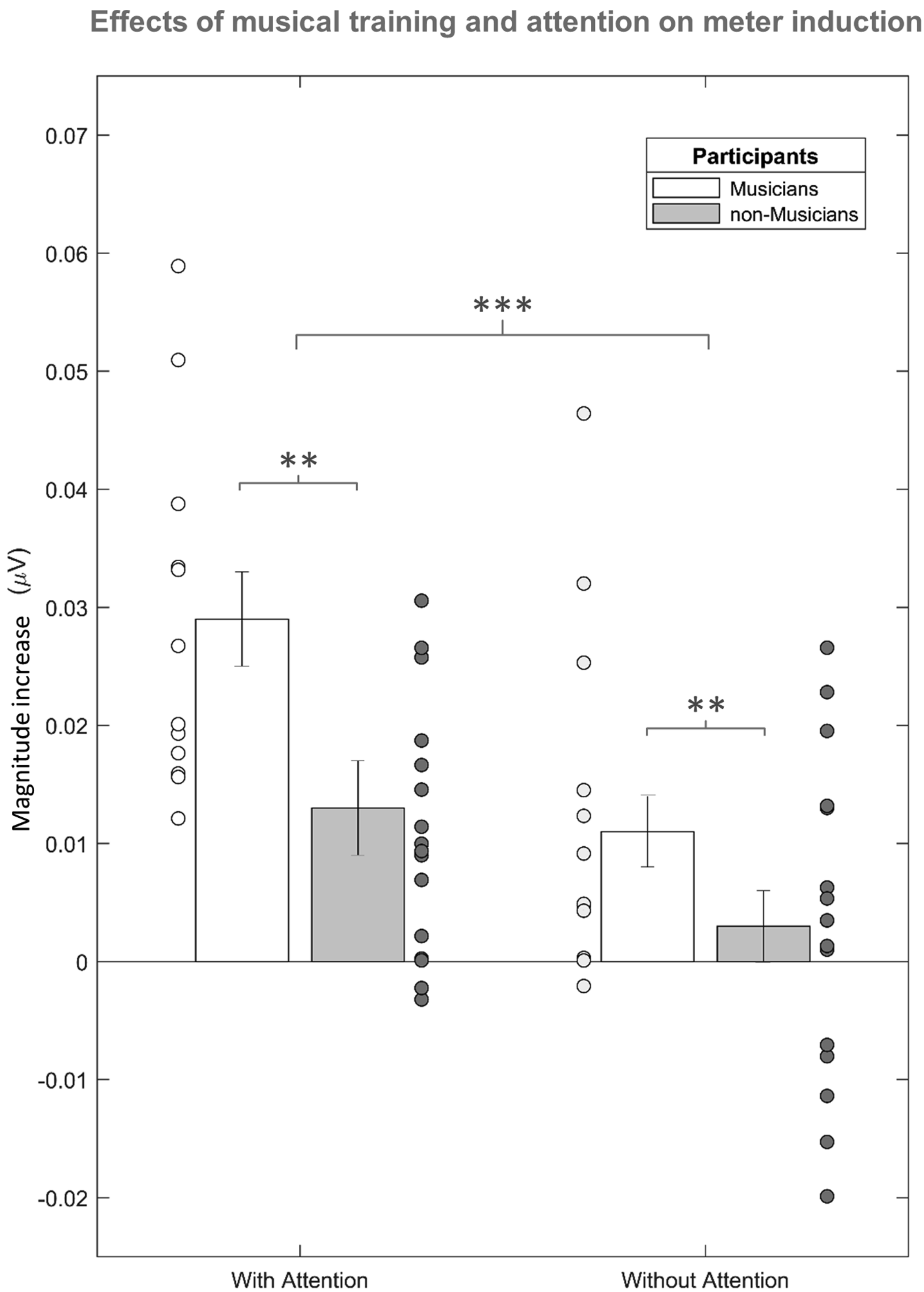

Fig. 7. Amplitude differences of musicians and non-musicians depending on attention. The six frequencies of interest and the three Spatial conditions are averaged at each bar. Their standard error bars are depicted. The pairwise comparison revealed that musicians had higher increases than non-musicians. The increases were also higher when the auditory stimuli were attended. The individual means were plotted next to the bars. The significance level is marked with asterisks: ** indicates $p < .01$ and *** indicates $p < .001$.

that the enhanced binaural hearing of the musicians (Parbery-Clark et al., 2013) may have helped them to entrain to the slow meter, while the lacking of musical training and the visual distractor may have hampered this slow entrainment in non-musicians despite their ability to entrain to the higher harmonic frequencies of the meter. It is also possible that the fast Fourier transform is reflecting the *quasi*-frequency-locked neural oscillations to the meter as higher peaks at their more synchronized harmonics.

Collapsing the amplitudes of musicians and non-musicians together, we observed that some peaks at the frequencies of the meter (0.8, 1.6, 3.2 and 4 Hz) were higher than in the control. Remarkably, only the Spatial 90° condition was different to the control at 0.8 Hz. This suggests that low level features of the sound are not enough to process and organize the isochronous sounds based on the auditory spatial contrasts in the intermediate conditions (Spatial 30° and 60°). The fact that participants were paying attention to a visual distractor may have disrupted the metrical processing of the beat in these conditions. Interestingly, the complete alternation of the sounds in the Spatial 90° condition may have overcome the scarcity of attentional resources. We highlight that musicians detected the ternary meter across all Spatial conditions. This suggests they may have had involuntary neural entrainment to it. Similarly, non-musicians were also aware of the cyclicity of the ternary pattern. This might be linked to the small peaks observed at meter-related frequencies. Although we asked participants to not attend the auditory stimuli, they may have found hard to avoid the rhythmic engagement, especially in musicians, who had an extended training in processing rhythms.

When we compared the modulation of neural entrainment across experiments, we observed higher peak increases in the spatial conditions for the attended auditory stimuli than for the unattended ones. This suggests that attention facilitates the processing of meter when the organization of the events depends on spatial cues. The comparison also revealed that musicians had an advantage in grouping the sounds in a ternary structure, regardless of the attentional demands. Finally, the *post-hoc* tests also revealed that the increases at meter-related frequencies were generally greater than those at beat-related frequencies. This directly reflects the contribution of metrical processing on the spatially-located sounds. In sum, the comparison across experiments made evident the role of attention and musical training in processing the beat within a metrical structure.

4. General discussion

Our study investigates how the brain interacts with alternating sounds to induce a metrical structure upon the beat. In Experiment 1, we observed that the spatial location of sounds modulated the rhythmic organization of the beat and triggered neural entrainment to ternary meter. Besides, formal training in music increased the amplitudes related to the SSEPs of the beat and the ternary meter. This does not necessarily mean that musicians have a better perception of the sound. Rather, musicians could be more accurate in entraining to the beat and the meter due to their long experience with the rhythmic dimension of music. In Experiment 2, we observed fewer significant differences between the Control and the Spatial conditions. Neural entrainment to the frequency of meter was only found in musicians. When we compared both experiments, we found that attention to the auditory stimuli increased the neural entrainment to the frequencies of the beat and the meter. This may indicate that the metrical organization of an alternating beat does not only depend on bottom-up processes of low-level cues, but also on top-down attentional mechanisms that help to process the spatial contrasts of the sounds as salient cues.

The present study suggests some differences between musicians and non-musicians regarding the neural processing of beat and meter using spatial sounds. Other evidence for these differences is found in more general sensorimotor synchronization tasks (for an extended review see Repp, 2005; Repp & Su, 2013) and brain responses measured by EEG (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Geiser, Sandmann, Jäncke, & Meyer, 2010). For instance, when tapping to music excerpts, musicians are found to access more easily a greater range of hierarchical levels to impose meter on the beat (Drake, Penel, & Bigand, 2000). The influence of music expertise on the processing of hierarchical temporal structures has recently been tested using EEG and frequency analyses (Stupacher, Wood, & Witte, 2017). The neural processing of complex rhythms (i.e. polyrhythms) was examined in musicians and non-musicians, and no differences appeared among them during the listening of superimposed meters. The differences only appeared when the participants maintained the metrically-organized beat silently after listening to the complex auditory stimuli.

In Experiment 1, musicians showed greater amplitudes than non-musicians at all the frequencies of interest: the beat (2.4 Hz), the meter (0.8 Hz) and their harmonics (1.6, 3.2, 4.0 and 4.8 Hz). The enhanced neural entrainment of musicians may be related to the familiarity with rhythmic patterns and the accuracy in processing them. In contrast, Stupacher et al. (2017) found that complexly arranged acoustic events are not processed differently by musicians and non-musicians. Since our stimuli are simpler, we could hypothesize that formal musical training might mainly affect the metrical schema underlying rhythms, but not the more complex patterns found in actual melodies. If this is true, then musicians may just have an advantage to apply a metrical schema from the moment they detect a repetitive pattern. In fact, the non-significant interaction between Group and Condition of Experiment 1 seems to support this idea: the amplitudes of musicians and non-musicians only differed in the Spatial 30° and Spatial 60° conditions, but not in the Spatial 90° condition, when the angular distance was the greatest. This suggests that musicians readily applied the ternary meter schema since the first Spatial condition, whereas non-musicians followed a more gradual entrainment to the ternary meter. Non-musicians may have needed more evident cues to engage in a metrical organization of the sounds. An alternative explanation could be that musicians have a better low-level auditory processing compared to non-musicians, rather than a better representation of the metrical schema per se. This should be further studied with comparative studies and neuroimaging techniques.

In Experiment 2, we did not find significant differences between musicians and non-musicians when we compared all the amplitudes across frequencies and conditions. However, we found group differences when we compared the increases of the peaks across experiments. In Experiment 2, only the Spatial 90° condition was different from the control at the frequency of the meter (0.8 Hz). The lack of SSEPs at the frequency of the meter in the spatial 30° and 60° conditions may indicate that spatial cues were omitted as reliable signals for grouping the sounds in a ternary pattern until they became completely evident in the Spatial 90° condition. While non-musicians did not show any trace of meter induction at 0.8 Hz, musicians had peaks at 0.8 Hz in the Spatial conditions, which could indicate that they were already projecting a metrical schema in an unconscious manner. In contrast, the absence of clear peaks at 0.8 Hz in non-musicians supports the idea that the metrical organization of the beat requires a top-down attentional mechanism rather than just a bottom-up pre-attentive auditory mechanism. This fits well with the proposal of primitive (bottom-up) and knowledge-based (top-down) components for Auditory Scene Analysis (Bregman, 1994). The former consists of a pre-attentive automatic parsing of the input based on basic auditory cues, while the latter consists of an endogenous process involving conscious attention and past experience with the sounds to organize the sensory information (Moore & Gockel, 2012; Trainor, 2015). Our findings are also in line with the study by Chapin et al. (2010), who found that attention to the auditory stimuli, rather than to a visual distractor, increased BOLD activity in areas related to beat and meter. Our second experiment therefore provides some evidence for the role of attention in sequential grouping. It also suggests that the background experience of musicians in processing rhythmic sounds could surpass the attentional demands imposed by the metrical grouping over spatial cues.

In our study, sounds were presented in sequences of alternating sounds following a ternary metrical structure: 1 sound on one side followed by 2 sounds on the opposite side. It is thus possible that the contralateral sound was perceived as a deviant stimulus within the sequence and could thus elicit early neural responses similar to the Mismatch Negativity (for details about MMNs, see Garrido, Kilner, Stephan, & Friston, 2009). If so, the periodicity of occurrence of these MMNs within the EEG signal may be reflected by a peak at the frequency of the meter. However, in Experiment 2 we did not observe a peak at 0.8 Hz for the contralateral sounds in non-musicians. The repetitive presentation of the contralateral sound was so predictable (24 times in each sequence) that the brain may have rapidly become habituated to the constant alternation elicited by the spatial contrast. Thus, the differences in the amplitudes related to the SSEPs of the beat and the ternary meter we observed in Experiment 1 seem to depend on the induction of the metrical structure.

Our experiments explicitly focus on the regular alternation of spatial sounds that occur at the frequencies of the musical beat and meter (i.e. delta-band). Previous studies focused on fast auditory changes in space (McAlpine, Haywood, Undurraga, & Marquardt, 2016; Nozaradan, Mouraux, & Cousineau, 2017; Undurraga, Haywood, Marquardt, & McAlpine, 2016) also using stimuli alternating from one side to the other, but their frequencies of interest were faster (from alpha-band to gamma-band) than the typical for beat and meter. While Undurraga et al. (2016) worked with a special binary structure elicited by six alternating fast sounds, Nozaradan et al. (2017) used quinary grouping, an alternation occurring every five periodic sounds. Contrastingly, our stimuli consist of another very typical metrical structure: the ternary grouping, which is very used in dances like the waltzes. Our findings therefore contribute to this line of research that aims to comprehend how changes in the acoustic space are processed at different timescales. The findings from Experiment 1, together with the previously mentioned studies, could be interpreted as if the automatic low-level processing of spatial contrasts does not depend on the rhythmic structure of the sounds. The findings from Experiment 2 point in another direction, as we did not observe neural entrainment at the frequency of meter in the Spatial 30° and 60° conditions. The results seem to indicate that attentional mechanisms may be a necessary requisite to process and establish the metrical organization of the beat. More research is however needed to conclude up to what point the contrasting spatial sounds are enough to trigger the higher cognitive structures of meter, either using other designs and distractors or applying distinct brain imaging methodologies.

The temporal orienting of attention has also been tested cross-modally. For instance, Bolger, Trost, and Schön (2013) found that the entrainment to meter affects the responses to both auditory and visual stimuli presented at distinct metrical positions. Furthermore, the neural entrainment to an imagined ternary meter projected on isochronous visual flashes supports the cross-modality of timing mechanisms (Celma-Miralles et al., 2016). While the previous study focuses on the temporal dimension of the visual modality, the present work focuses on the spatial dimension of the auditory modality. Both studies demonstrate neural entrainment across domains, suggesting that meter induction can be extended beyond audition to allow for the processing of rhythms made up with distinct perceptual cues. One can speculate that the domain-general areas activated in the beat-based and duration-based timing networks (Teki, Grube, & Griffiths, 2012), which involve premotor cortex and supplementary motor areas, the basal ganglia and the cerebellum (among others), open the door to integrate neural information beyond the auditory modality. In this line, Schubotz et al. (2000) reported that the neural network of time perception relies on the cerebral areas activated for temporal planning and movement coordination. Although we have not identified any motor region of the brain, our results support that the integration of spatial information as metrical cues can be observed in musicians and non-musicians. Thus, rhythmic processing might use information from different domains to build up metrical structures and organize the events in temporal patterns. Interestingly, this idea of a cross-modal hierarchical organization of temporal events in the brain aligns well with the dendrophilia hypothesis exposed by Fitch (2014), which proposes that humans are specially inclined to parse and store tree-like structures in music and language. In fact, both systems may share some of the syntactic resources underlying this combinatorial ability (Patel, 2003), but they may operate on distinct structural representations: words or tones.

The analysis of Experiment 1 exploring the relationship between the lateralization of the stimuli and the hemispheric lateralization of the electrodes in the scalp revealed that there are neither differences regarding the onset lateralization of the stimuli nor any effect of ipsilateral and contralateral electrodes. This implies that the counterbalanced presentation of the stimuli did not alter the results we observed and that there is no hemispheric specialization for spatially-induced meter perception. Gutschalk and Steinmann (2015) found a dominant contralateral processing of a monaural noise in the auditory cortex and subcortical regions using fMRI data. They also observed bilateral processing when some aspects of the sound, such as the amplitude, are slowly modulated. However, they did not find any consistent contralateral effect when using MEG data, likely due to the steady-state activity phase elicited by the amplitude modulation. Similarly, our recordings of neural entrainment did not show any difference between right or left regions of the scalp for the SS-EPs related to the beat and meter. These results may indicate that our task does not clearly elicit any lateralized processing of meter for the spatial sounds. Another possibility for the lack of lateralization in the neural entrainment is that the right hemispheric bias (Brunetti et al., 2005; Johnson & Hautus, 2010; Kaiser et al., 2000; Zatorre & Penhune, 2001) for sound localization could have interacted with a lateralized processing of meter that masked any clear effect between the recordings of the right and left electrodes. Further analyses on source localization or comparing the topographies at each frequency of interest across conditions could possible reveal lateralization effects in meter induction on spatial sounds.

The differences between Experiment 1 and Experiment 2 support the idea that beat-related and meter-related steady-state evoked potentials are not just a reflection of the stimulus proprieties, i.e. sensory entrainment, but also a reflection of endogenous processes that modulate attention and generate predictions for relevant points in time (i.e. neural entrainment; see Bauer, Jaeger, Thorne, Bendixen, & Debener, 2015; Large, Herrera, & Velasco, 2015; Nozaradan, Peretz, & Keller, 2016; Stupacher, Witte, Hove, & Wood, 2016). This modulation of attention over time was first proposed by Jones (1976) and subsequently studied as the dynamic attending theory (Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999). We observed that putting attention on the alternating sounds was important for the neural entrainment to the ternary meter defined over spatial cues. Similarly, attention to complex rhythms was found to activate the areas of the brain related to beat and meter (Chapin et al., 2010). Taken altogether, these findings seem to indicate that attentional mechanisms are necessary to establish a consistent metrical organization of the beat.

Our findings open the door to further studies exploring how the sounds are processed in a metrical structure elicited by spatial cues. For example, it would be very revealing to test patients with auditory neglect on a similar task. These patients usually present inattention to stimuli within the right or left hemifield due to unilateral hemispheric lesions. For instance, Zatorre and Penhune (2001) tested how patients with unilateral temporal lobe excisions dealt with spatially located clicks and found that, despite individual differences, the auditory spatial processes regarding source identification seem to rely more on the right superior temporal cortex for both spatial hemifields. The study of a similar population could clarify whether there is a lateralized metrical processing in the brain and, if this were the case, how it deals with acoustic cues that only vary on their spatial localization. Another way to disentangle the effect of the located sound on the metrical structure would be by scattering the alternation of the sound sources without following any metrical structure. Then, we could ask participants to just listen or to internally project a ternary meter on the randomly-located sounds. This study may provide more evidence for similitudes and differences between the top-down projection of meter and the processing of randomly presented spatial sounds.

5. Conclusions

The present work focuses on the relationship between the spatial dimension of sounds and the processing of rhythms. We used sounds that cyclically alternated at different angles to elicit neural entrainment to the frequency of a ternary structure, with its downbeat at 0.8 Hz. We compared the EEG recordings of spatially alternating sounds to those of a control with a steady beat at 2.4 Hz. When attention was placed on the auditory stimuli, we observed neural entrainment to the meter and higher amplitudes for musicians compared to non-musicians. When attention was put on the visual distractor, the neural entrainment to the meter was inconsistent and the effects of formal training in music were hidden. The lack of entrainment to meter in the Spatial 30° and 60° conditions of experiment 2 suggests that the SSEPs observed in Experiment 1 represent more than just low-level auditory responses. Attention may have played a role in facilitating the metrical organization of the beat using the spatial contrasts as relevant cues. Finally, the comparison of both experiments showed that musicians had greater increases of the magnitudes than non-musicians in the Spatial conditions, both at meter-related and beat-related frequencies. This reinforces the idea that the neural processing of beat and meter is facilitated by the previous experience with rhythmic events during the long periods of formal music training.

Acknowledgements

This work was supported by the European Research Council (ERC) Starting Grant agreement n. 312519 and by the Spanish Ministerio de Economía y Competitividad (MEC) FPI grant BES-2014-070547. We would like to thank Laia Garcia Vernet for her help during the design of the stimuli, as well as the reviewers and colleagues from the CBC, the Z-lab and the MIB for all their insightful comments.

Declaration of Competing Interest

We declare no competing interests.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandc.2019.103594.

References

Bauer, A. K. R., Jaeger, M., Thorne, J. D., Bendixen, A., & Debener, S. (2015). The auditory dynamic attending theory revisited: A closer look at the pitch comparison task. *Brain Research, 1626*, 198–210.
Bennett, J. C., Barker, K., & Edeko, F. O. (1985). A new approach to the assessment of stereophonic sound system performance. *Journal of the Audio Engineering Society, 33*(5), 314–321.
Bishop, L., & Goebl, W. (2017). Beating time: How ensemble musicians' cueing gestures communicate beat position and tempo. *Psychology of Music* 0305735617702971.
Bolger, D., Trost, W., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica, 142*(2), 238–244.
Borgo, M., Soranzo, A., & Grassi, M. (2012). *MATLAB for psychologists.* Springer Science & Business Media.
Born, G. (2013). *Music, sound and space: Transformations of public and private experience.* Cambridge University Press.
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound.* MIT press.
Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14*(4), 362–366.
Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Della Penna, S., Ferretti, A., ... Romani, G. L. (2005). Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. *Human Brain Mapping, 26*(4), 251–261.
Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience, 8*.
Butler, R. A., & Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *The Journal of the Acoustical Society of America, 61*(5), 1264–1269.
Celma-Miralles, A., de Menezes, R. F., & Toro, J. M. (2016). Look at the beat, feel the meter: Top–down effects of meter induction on auditory and visual modalities. *Frontiers in Human Neuroscience, 10*.
Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: The role of attending. *Frontiers in Psychology, 1*, 224.
Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854.
Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception: An Interdisciplinary Journal, 18*(1), 1–23.
Feddersen, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (1957). Localization of high frequency tones. *The Journal of the Acoustical Society of America, 29*(9), 988–991.
Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience, 7*.
Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews, 11*(3), 329–364.
Fujiki, N., Riederer, K. A., Jousmäki, V., Mäkelä, J. P., & Hari, R. (2002). Human cortical representation of virtual auditory space: Differences between sound azimuth and elevation. *European Journal of Neuroscience, 16*(11), 2207–2213.
Fujioka, T., Zendel, B. R., & Ross, B. (2010). Endogenous neuromagnetic activity for mental hierarchy of timing. *Journal of Neuroscience, 30*(9), 3458–3466.
Garcia-Vernet, L. (2017). Object-based cross-modal enhancement in sound localization (Thesis, Universitat Pompeu Fabra, Barcelona, Spain). Retrieved from https://repositori.upf.edu/handle/10230/30812.
Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology, 120*(3), 453–463.
Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception–training increases hierarchical metre processing. *European Journal of Neuroscience, 32*(11), 1979–1985.
Gelfand, S. (1982). Hearing: An introduction to psychological and physiological acoustics. *Journal of Neurology, Neurosurgery, And Psychiatry, 45*. https://doi.org/10.1136/jnnp.45.12.1175-b.
Gilkey, R., & Anderson, T. R. (2014). *Binaural and spatial hearing in real and virtual environments.* Psychology Press.
Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience, 19*(5), 893–906.
Gutschalk, A., & Steinmann, I. (2015). Stimulus dependence of contralateral dominance in human auditory cortex. *Human Brain Mapping, 36*(3), 883–896.
Harley, M. A. (2016). *Space and spatialization in contemporary music: History and analysis, ideas and implementations.* Lulu Press.
Hong, S. (2015). Neural correlates of beat and metre perception: The role of the inferior frontal gyrus (Doctoral dissertation, University of Edinburgh, Edinburgh, United Kingdom). Retrieved from https://www.era.lib.ed.ac.uk/bitstream/handle/1842/21112/Hong2015.pdf.
Johnson, B. W., & Hautus, M. J. (2010). Processing of binaural spatial information in human auditory cortex: Neuromagnetic responses to interaural timing and level differences. *Neuropsychologia, 48*(9), 2610–2619.
Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323.
Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313–319.
Kaiser, J., Lutzenberger, W., Preissl, H., Ackermann, H., & Birbaumer, N. (2000). Right-hemisphere dominance for the processing of sound-source lateralization. *Journal of Neuroscience, 20*(17), 6631–6639.
Kuck, H., Grossbach, M., Bangert, M., & Altenmüller, E. (2003). Brain processing of meter and rhythm in music. *Annals of the New York Academy of Sciences, 999*(1), 244–253.
Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119.
Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience, 9*, 159.
Lee, K. M., Barrett, K. C., Kim, Y., Lim, Y., & Lee, K. (2015). Dance and music in "gangnam style": How dance observation affects meter perception. *PloS One, 10*(8), e0134725.
Lerdahl, F., & Jackendoff, R. (1981). On the theory of grouping and meter. *The Musical Quarterly, 67*, 479–506.
London, J. (2012). Three things linguists need to know about rhythm and time in music.
Luck, G., & Sloboda, J. A. (2007). An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Performance Research, 1*, 26–46.
McAlpine, D., Haywood, N., Undurraga, J., & Marquardt, T. (2016). Objective measures of neural processing of interaural time differences. *Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp. 197–205). Cham: Springer.
Moore, B. C., & Gockel, H. E. (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal Society B, 367*(1591), 919–931.
Moroney, D. (2007). Alessandro Striggio's mass in forty and sixty parts. *Journal of the American Musicological Society, 60*(1), 1–70.
Münte, T. F., Kohlmetz, C., Nager, W., & Altenmüller, E. (2001). Neuroperception: Superior auditory spatial tuning in conductors. *Nature, 409*, 580.
Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münte, T. F. (2003). The fate of sounds in conductors' brains: An ERP study. *Cognitive Brain Research, 17*, 83–93.
Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658), 20130393.
Nozaradan, S. (2017). Musical rhythm embedded in the brain: Bridging music, neuroscience, and empirical aesthetics. In Enhancing participation in the arts in the EU (pp. 99-113). Cham: Springer.
Nozaradan, S., Mouraux, A., & Cousineau, M. (2017). Frequency tagging to track the neural processing of contrast in fast, continuous sound sequences. *Journal of Neurophysiology, 118*(1), 243–253.
Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual differences in rhythmic cortical entrainment correlate with predictive behavior 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*(28), 10234–10240.
Pachet, F., & Delerue, O. (1998). A mixed 2d/3d interface for music spatialization. In Virtual worlds (pp. 298–307). Berlin/Heidelberg: Springer.
Parbery-Clark, A., Strait, D. L., Hittner, E., & Kraus, N. (2013). Musical training enhances neural processing of binaural sounds. *Journal of Neuroscience, 33*, 16741–16747.
Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674.
Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., ... Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain: A Journal of Neurology, 120*(2), 229–243.
Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour, 1*, 0007.
Ravignani, A., Thompson, B., Grossi, T., Delgado, T., & Kirby, S. (2018). Evolving building blocks of rhythm: How human cognition creates music via cultural transmission. Annals of the New York Academy of Sciences.

Rayleigh, L. (1907). XII. On our perception of sound direction. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 13(74), 214–232.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969–992.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403–452.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., & Nielsen, M. (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience, 19*(22), 10074–10081.

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences, 112*(29), 8987–8992.

Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. In Progress in brain research (Vol. 217, pp. 37–55). Elsevier.

Schubotz, R. I., Friederici, A. D., & Von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *Neuroimage, 11*(1), 1–12.

Stupacher, J., Witte, M., Hove, M. J., & Wood, G. (2016). Neural entrainment in drum rhythms with silent breaks: Evidence from steady-state evoked and event-related potentials. *Journal of Cognitive Neuroscience, 28*(12), 1865–1877.

Stupacher, J., Wood, G., & Witte, M. (2017). Neural entrainment to polyrhythms: A comparison of musicians and non-musicians. *Frontiers in Neuroscience, 11*, 208.

Teki, S., Grube, M., & Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience, 5*, 90.

Tervaniemi, M. (2009). Musicians-same or different? *Annals of the New York Academy of Sciences, 1169*, 151–156.

Tervaniemi, M., Castaneda, A., Knoll, M., & Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: Event-related potential and behavioral indices. *Neuroreport, 17*, 1225–1228.

Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience, 30*, 1636–1642.

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society B, 370*(1664), 20140089.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception: An Interdisciplinary Journal, 28*(1), 59–70.

Undurraga, J. A., Haywood, N. R., Marquardt, T., & McAlpine, D. (2016). Neural representation of interaural time differences in humans—An objective measure that matches behavioural performance. *Journal of the Association for Research in Otolaryngology, 17*(6), 591–607.

Ungan, P., Yagcioglu, S., & Goksoy, C. (2001). Differences between the N1 waves of the responses to interaural time and intensity disparities: Scalp topography and dipole sources. *Clinical Neurophysiology, 112*(3), 485–498.

Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., & Tervaniemi, M. (2012). The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia, 50*, 1432–1443.

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences, 13*, 532–540.

Zatorre, R. J., & Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex. *Journal of Neuroscience, 21*(16), 6321–6328.