"Look at the beat, feel the meter"

Top-down effects of meter induction on auditory and visual modalities

by

Alexandre Celma Miralles

and supervised by

Dr. Juan Manuel Toro Soto

Thesis submitted in partial fulfilment of the requirements for the degree

Master in Brain and Cognition

September, 2015

Center for Brain and Cognition

Department of Communication and Information Technologies

Universitat Pompeu Fabra

Aquesta tesina no hauria estat possible sense

el recolzament, pacient i atent del meu supervisor ,

els recíprocs ànims de les companyes i companys de màster,

ni els consells i ajuda de l'equip investigador i tècnic del CBC.

Encara menys hauria estat possible sense

el caliu constant de familiars i amistats de sempre,

ni l'ajuda ferma, detallista i incondicional, de la meva parella.

Fuquay-Varina (USA), 16th August 2015

ABSTRACT

This study seeks to shed light on the nature of meter induction, the musical ability to organize isochronously-perceived beats in hierarchical structures. Recent research demonstrated top-down effects on meter induction in the auditory modality. We aim to assess whether those effects are also present in the visual domain. Sixteen musicians were asked to internally project binary (i.e. a strong-weak pattern) and ternary (i.e. a strong-weak-weak pattern) meter onto separate, analogous visual and auditory isochronous stimuli. Participants were presented with sequences of tones or blinking circular shapes (i.e. flashes) at 2.4 Hz while their electrophysiological responses were recorded. The frequency analysis of the elicited steady-state evoked potentials allowed us to compare the frequencies of the beat (2.4 Hz), its first harmonic (4.8 Hz), the binary subharmonic (1.2 Hz), and ternary subharmonic (0.8 Hz) within and across modalities. Taking the amplitude spectra into account, we observed an effect at 0.8 Hz in the ternary condition for both modalities, which suggests a cross-modal meter induction. There was also an interaction between modality and magnitude at 2.4 and 4.8 Hz. Looking at the power spectra, we observed significant differences from zero in the auditory, but not in the visual, binary condition at 1.2 Hz. Finally, a generalized linear model revealed an interaction between dance training and the visual modality. These findings suggest that meter processing is modulated by top-down mechanisms that interact with our perception of rhythmic events and that such modulation can also be found in the visual domain.

RESUM

Aquest estudi pretén discernir la naturalesa de la inducció mètrica, l'habilitat musical d'organitzar la pulsació percebuda isocrònicament en estructures jeràrquiques. Recentment s'han demostrat efectes *top-down* de la inducció mètrica per a la modalitat auditiva. Volem saber si els mateixos efectes apareixen per a la modalitat visual. Vam demanar a setze músics que projectessin internament un metre binari (fort-feble) i un de ternari (fort-feble-feble) sobre estímuls auditius i visuals isocrònics i anàlegs. Es van presentar seqüències isocròniques de tons o de cercles parpellejant (flaixos) a 2.4 Hz mentre es gravaven les respostes electrofisiològiques. L'anàlisi de freqüències dels potencials evocats en estat estacionari (SS-EP en anglès) permeten de comparar les freqüències de la pulsació (2.4 Hz), el seu primer harmònic (4.8 Hz), el subharmònic binari (1.2 Hz) i el subharmònic ternari (0.8 Hz) dins de cada modalitat i entre elles. Amb els espectres d'amplitud s'observa un efecte a 0.8 Hz a la condició ternària dins les dues modalitats, cosa que suggereix una inducció mètrica intermodal. Apareix també una interacció entre modalitat i magnitud a 2.4 i 4.8 Hz. A través de l'espectre de potència, s'observa per a la condició binària auditiva un valor significativament diferent de zero a 1.2 Hz. Finalment, el model lineal generalitzat revela una interacció entre dansa i la modalitat visual. Aquests resultats suggereixen que el processament del metre es modula per mecanismes *top-down* que interaccionen amb la percepció de ritmes, i que aquesta modulació també ocorre en el domini visual.

RESUMEN

Este estudio pretende discernir la naturaleza de la inducción métrica, la habilidad musical de organizar la pulsación percibida isocrónicamente en estructuras jerárquicas. Recientemente se han demostrado efectos *top-down* de la inducción métrica para la modalidad auditiva. Queremos saber si los mismos efectos aparecen para la modalidad visual. Pedimos a dieciséis músicos que proyectaran internamente un metro binario (fuerte-débil) y ternario (fuerte-débil-débil) sobre estímulos auditivos y visuales isocrónicos y análogos. Se presentaron secuencias de tonos y de círculos parpadeando (flashes) a 2.4 Hz, mientras se grababan las respuestas electrofisiológicas. El análisis de frecuencias de los potenciales evocados en estado estacionario (SS-EP en inglés) permiten comparar las frecuencias de la pulsación (2.4 Hz), su primer armónico (4.8 Hz), el sub-armónico binario (1.2 Hz) y el sub-armónico ternario (0.8 Hz) dentro de cada modalidad y entre ellas. Con los espectros de amplitud se observó un efecto a 0.8 Hz a la condición ternaria dentro de las dos modalidades, lo que sugiere una inducción métrica intermodal. Aparece también una interacción entre modalidad y magnitud a 2.4 y 4.8 Hz. A través del espectro de potencia, se observó un valor significativamente diferente de cero a 1.2 Hz para la condición binaria auditiva. Finalmente, el modelo lineal generalizado reveló una interacción entre danza y la modalidad visual. Estos resultados sugieren que el procesamiento del metro se modula por mecanismos *top-down* que interaccionan con la percepción de ritmos, y que esta modulación también ocurre en el dominio visual.

TABLE OF CONTENTS

I. INTRODUCTION

1. *Music and meter*

The faculty of music, similar to language, is a universal innate capacity found cross-culturally around the world. Music consists of an organized arrangement of sounds and silences that evoke emotions and involves people in a social interactive performance, made up of gestures, sounds, and shared intentions and moods. Music encodes tension-relaxation patterns (Lerdahl and Jackendoff, 1983; Jackendoff and Lerdahl, 2006; Lerdahl and Krumhansl, 2007) within affective-gestural and socio-intentional systems.

Music mainly deals with temporal structures, organized through meter, grouping, and tonal encoding of pitch. These properties allow us to predict, plan and process upcoming events over time. Rhythms are perceptual events placed over time, while tones are perceptual events placed over pitch frequency. While the former involves beat and meter, the latter involves musical notes and tonality. The hierarchical organization of beats leads to meter, while the hierarchical organization of musical notes leads to tonality.[1] In the present study, we will focus on meter, the hierarchical organization of periodic beats in sequences showing strong and weak patterns, where the downbeat usually occurs periodically at a subharmonic frequency of the beat. This mechanism is claimed to be unique to humans (Fitch, 2013).

Meter can be characterized by three features[2]: (i) it is totally distinct from rhythmic grouping, (ii) it is a top-down mechanism based on prior knowledge that influences the active process of rhythmic perception and (iii) musical timing shows distinct features at different timescales. Regarding the first point, it has been claimed that meter is "an endogenous sense of hierarchically arranged levels of pulsation", having a salient level, *the tactus*, which tends to induce periodic body movements (Keller, 2012). This is a strongly predictive mechanism which informs *when* an event will occur over time, thanks to learned statistical schemes of regularities. According to the second point, rhythm perception is active because top-down processes such as "attentional resource allocation and mental schemas" lead to a subjective metricization (Keller, 2012), that is, the perception of regular alternations between strong and weak beats. Regarding the third point, there are different levels of music structure that can be computed. Research has demonstrated that our dynamic neural oscillations synchronize at different frequency bands depending on the level of music structure they are computing. In

[1] In addition, another hierarchical structure, *grouping*, works on a macroscopic layer, making use of rhythm and pitches (London, 2012, Keller, 2012).
[2] Presented by Justin London (2012) and complemented by Katie Overy (2012) and Peter E. Keller (2012).

fact, experiments by Nozaradan and collaborators (2011) have revealed that the musical *tactus* is supported by one of these time-scales, combining beats into unitary chunks.

2. The frequency-tagging of beat and meter

Several studies have explored the dynamic neural oscillations linked to meter induction. Most of them use a frequency-tagging approach. The frequency-tagging approach is an electrophysiological method based on the "periodic change in voltage amplitude in the electrical activity recorded on the human scalp by EEG" elicited by a periodic property of a stimulus (Nozaradan, 2014). Since these neural responses are stable in phase and amplitude over time, they are called steady-state evoked potentials (SS-EP). Instead of looking at the time domain, SS-EPs allow for analysis of the frequency domain, revealing narrow-band peaks at frequencies related to periodic properties of the stimulus. Although multiple peaks appear in the frequency spectrum induced by the rhythmic pattern envelope, the responses elicited at the beat and meter frequencies are selectively enhanced. One idea that has received empirical support from this approach is that perceived periodicities of beat and meter are mental constructs (i.e. abstracted percepts); natural rhythms are neither totally periodic, nor does meter need to be perceptually marked exogenously through pitch, amplitude, or timbre modulations.

One advantage of the frequency-tagging approach is that neural responses at the frequencies of the expected beat and meter can be objectively identified and quantified through frequency domain analyses without explicit behavioral responses biasing the cognitive measures. The emergence of such neural responses could be due to the existence of a network of non-linear oscillators (Nozaradan, 2014). When looking at the beat and meter induced SS-EPs, this method reveals a "mechanism through which attentional and perceptual processes are dynamically modulated as a function of time" (Nozaradan, 2014). Crucially, top-down effects also are captured by this method. For instance, when an internally driven meter that is not present in the stimulus is imposed on the stimulus, this leads to a SS-EP at the target frequency (Nozaradan et al., 2011). This method is fundamental for our study in order to investigate mentally induced meter in the visual modality.

Thus, electrophysiological correlates of beat and meter induction can be found experimentally (Nozaradan, 2013, 2014). In the study by Nozaradan et al. (2011), eight participants were first asked to listen to a pseudo-periodic auditory stimulus and then to project a binary and ternary meter structure on it. The stimulus lasted for 33 seconds, and its amplitude was modulated at a rate of 2.4 Hz and 11 Hz. After obtaining the SS-EPs for each participant in each condition, frequency analyses revealed significantly different magnitudes at the beat frequency and its subharmonics $f/2$ and $f/3$. The authors found specific meter-induced neural entrainments at 1.2 Hz for the binary condition and at 0.8 Hz for the ternary condition. With these results, Nozaradan and collaborators claim that the voluntarily imposition of a specific frequency and phase on a stimulus modifies its metrical interpretation. That is, they demonstrated that meter can be induced by top-down processes. Moreover, these metrical periodicities are reflected in the brain through the increase of neural activity at the same frequencies and thus detectable via EEG.

Beat and meter directly relate to periodicities in the temporal structure of music and dance, and its perception often co-occurs with visual movements, as is the case when following a chamber music partner, watching a conductor directing an orchestra, or dancing in synchronization. The range of the frequencies for beat and meter appears to be restricted between 0.5 and 5 Hz; in other words, humans feel comfortable with an inter-onset interval (IOI), the time between the onset of a perceptual event and the following onset, of between 2 to 0.2 seconds (Repp, 2005). This range may also be optimal for prediction and coordination of movements. In fact, it has been hypothesized that the tight coupling of sensory and motor neural processes that coordinate the dynamics of distant brain areas is what permits rhythmic sensorimotor synchronization (Nozaradan et al., 2015). This is the reason why humans spontaneously entrains movements to the beat or its harmonics (Toiviainen et al., 2010; Leman and Naveda, 2010), as well as why motor actions affects the way beat and meter are perceived (Chemin et al., 2014). However, to date no study has directly tested meter induction in the visual modality without the involvement of motor acts (Merchant et al., 2015).

3. Our hypothesis

In the present study, we explore meter induction in the visual modality. We extend the findings on meter reported in Nozaradan et al. (2011) in the acoustic modality and predict correlations between the auditory and the visual domain. Some experiments based on beat tapping have explored meter induction in the visual and in the tactile domain (see Repp, 2005). However, to the best of our knowledge, no study has explored beat and meter in the visual domain without the involvement of motor acts. We approach this issue by asking some musicians to project a binary and ternary meter on analogous visual and auditory isochronous stimuli while their electroencephalographic activity was recorded. Our objective is to investigate whether meter, the hierarchical organization of beat in strong-weak patterns, is domain-specific in humans or applicable to other modalities, specifically vision.

II. MATERIALS AND METHODS

1. Participants

Sixteen healthy musicians were included in the present study (9 females, 6 left-handers, mean age: 23.375 ± 3.845, age range: 18 to 35). There were 17 participants in total, but one male was excluded due to an excess of artifacts in the EEG data. All participants were selected from a prior questionnaire and signed a written consent form. We chose sixteen participants in order to counterbalance all the possible blocks and conditions of the study. We tested musicians since they have to deal with beat and meter more often than non-musicians. They all had extensive musical experience, starting at 6.31 ±2.3 years of age. Six females reported some training in dance: three had spent between 6 and 11 years in dance school, while the other 3 had less than two years of experience. No participant reported any history of hearing, visual, motor, or psychiatric disorders, and all participants had normal or corrected-to-normal vision.

2. Stimuli

The auditory and visual stimuli were created using Matlab (v.2013, The MathWorks) and its Psychtoolbox extension. Our stimuli consisted of isochronous sequences of either tones (in the auditory condition) or flashes (in the visual condition). They were presented at a frequency of 2.4 Hz (IOI = 416.66 ms). Every sequence lasted for 35 seconds and comprised 84 tones or flashes. This frequency was chosen because all of our target frequencies fall within the ecological range of tempo perception and production (Nozaradan et al., 2011; Vialatte et al., 2009). Since our aim is to compare periodic beats in two different modalities, we did not create a pseudo-periodic stream as in Nozaradan et al. (2011), which could have made the results harder to interpret. Every stimulus progressively diminishes until the next one appears, thus marking the onset of the beat with the maximum intensity of sound (in the auditory condition) or with the maximum intensity of light (in the visual condition).

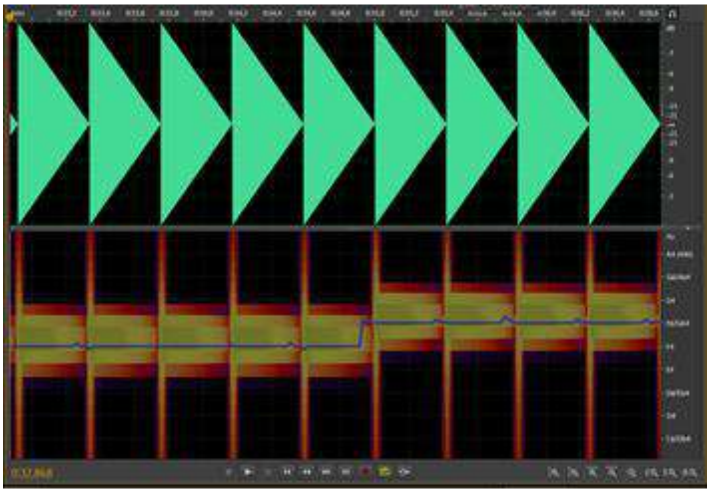


Figure 1. Sound envelope of the auditory stimuli, modulated at a rate of 2.4 Hz. The pitch was raised after every 35-second sequence.

Each condition consisted of eight auditory or visual 35-second sequences. After each 35-second sequence, the pitch or the color of the stimuli was changed in order to maintain participants' attention. No perceptual cues were used to differentiate each stimulus at a metrical level.

The auditory stimuli were presented through two speakers placed 70 cm in front of the participant at a comfortable hearing level. We converted a pure sinusoidal tone into stereo by using Audacity. These pure tones were raised half a tone in each 35-second sequence (see Figure 1), going from an F_4 up to a C_5. The whole auditory condition, therefore, consisted of eight different sequences of 84 sinusoidal tones: F_4 (349.2 Hz), $F\#_4$ (370.0 Hz), G_4 (392.0 Hz), $G\#_4$ (415.3 Hz), A_4 (440.0 Hz), $A\#_4$ (466.2 Hz), B_4 (493.9 Hz) and C_5 (523.3 Hz). Figure 1, taken from Adobe Audition CS6, shows the amplitude modulation of the stimuli over time.

The visual stimuli were presented on a computer screen placed 70 cm in front of the participant. A colored circle was placed at the center of the screen with a black background and had a radius of 45 mm. To create the blinking effect, we progressively diminished its luminescence until it turned completely black (see Figure 2). These flashes changed color after every 35-second sequence of 84 flashes with the following RGB progression: red (255 0 0), orange (255 128 0), yellow (255 255 0), green (128 255 0), turquoise (0 255 128), light blue (0 255 255), dark blue (0 0 255), and violet (128 0 255).

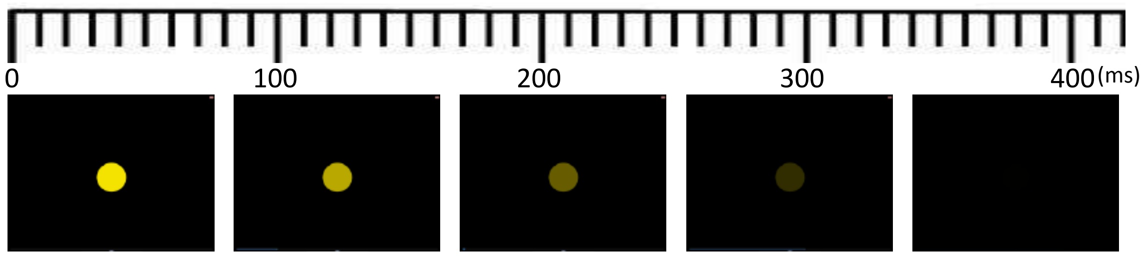


Figure 2. Visual stimuli over time (ms). These five frames were chosen among others because they perfectly depict the modulation of luminescence from maximum to minimum in 416.66 ms. The repetition of this modulation led to a flash effect tuned to 2.4 Hz. The color was changed in every 35-second sequence to maintain the participants' attention.

3. Procedure

The subjects were seated in a comfortable armchair in a soundproof room with the keyboard placed on their lap. Psychtoolbox (Matlab, The Mathworks) was used to run the experiment. Our study consisted of three conditions: the passive beat condition (which served as a control), the binary imagery task, and the ternary imagery task. These three conditions were the same for the auditory and visual blocks.

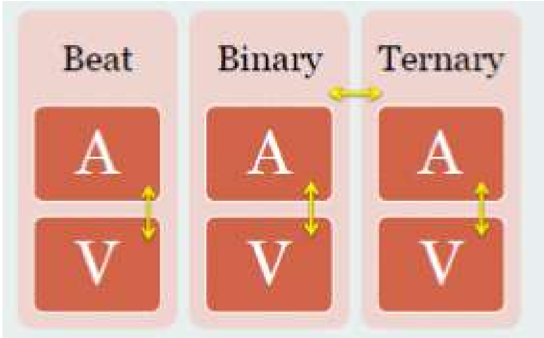


Figure 3. Sixteen counterbalanced conditions. The study has three conditions (beat, binary, and ternary meter) for each block (auditory [A] or visual [V] stimuli). The conditions were presented in a different order for each participant. The arrows represent the counterbalance of binary and ternary imagery tasks, as well as the auditory and visual modalities.

Every participant was presented with a different order of blocks and conditions, whereby the control condition was always presented first, and the visual and auditory blocks were alternated to form a total of 16 possible combinations in order to counterbalance the stimuli. During the control, the participants were either looking at the flashes or listening to the sounds passively. In order to avoid the induction of an involuntary binary meter in the control condition, the well-known *tick-tock effect*[3], we reminded the participants at the beginning of each 35-second sequence to perceive every beat individually, that is, as independent from the previous and the following tone or flash.

In contrast to the control condition, during the binary and ternary tasks, participants were asked to mentally project a binary structure (strong-weak pattern) or a ternary structure (strong-weak-weak pattern) on the same perceptual stimuli as presented in the control condition. In other words, they silently projected a metrical structure focusing on the subharmonics of the beat: $f/2$ (1.2 Hz) for the binary meter or $f/3$ (0.8 Hz) for the ternary meter. Participants were asked to start the meter imagery task as soon as the first stimuli was presented, and at the end of each 35-second sequence, they had to report whether the last beat was strong or weak as a way to maintain their attention and to make sure they were focused on the task. Because the participant had to press the space bar to go on to the next block, they were allowed to take a break when needed. The experimenter was out of the soundproof room during all the recordings but went in during the instructions between blocks (every two conditions) to solve any doubt the participant might have had about the meter imagery task. These moments were also used to remind the participant to neither move nor speak during the tasks. Once the study finished, a questionnaire was presented to the participant to evaluate and comment the stimuli and the difficulty of each task. Therefore, two questionnaires were requested for this study, before the study to verify the musical abilities of each participant, and just after the experiment to take into account their opinions and difficulties related to the experimental stimuli and their strategies to project meter.

[3] It consists of a subjective accentuation of identical sound events which makes isochronous sequences be perceived as unequal, showing a default pattern of binary metrical structure (Brochard et al., 2003).

4. EEG analyses

The EEG was recorded using an actiCAP with 60 electrodes placed on the scalp according to the International 10/10 system (Fp1, Fp2, AF7, AF3, AF4, AF8, F7, F3, F1, Fz, F2, F4, F8, FT9, FT7, FT8, FT10, FC5, FC3, FC1, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, T7, T8, TP9, TP7, TP8, TP10, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO9, PO3, POz, PO4, PO10, O1, Oz, O2). Vertical and horizontal eye movements were monitored using two electrodes placed on the infra-orbital ridge and the outer canthus of the right eye. Two additional electrodes placed on the left and right mastoid. The signals were referenced to the FCz channel and all electrode impedances were kept below 25 kΩ. The signals were amplified and digitized at the sampling rate of 1000 Hz.

Preprocessing of the continuous EEG recordings was implemented using BrainVision Analyzer 2.1 (Brain Products GmbH). First, any channel that appeared flat or noisy was interpolated from the surrounding channels via spherical spline interpolation. All the channels were then filtered using a zero-phase Butterworth filter to remove slow drifts in the recordings with a notch filter at 50 Hz, a high pass filter at 0.1Hz (48 dB/oct) and low pass filter at 10 Hz (time constant 1.591549, 48 dB/oct). Subsequently, eye blinks and muscular movements were corrected using the Ocular Correction ICA. Finally, the filtered EEG data was segmented into 35-second sequences that were grouped together according to each condition and block. These files were then converted into Matlab files, and all further statistical analyses were performed in Matlab v. 2013b (MathWorks) and SPSS (version 19, IBM).

5. Frequency analyses

For each condition and block, eight epochs lasting 32.5 seconds were obtained by removing the first 2.5 seconds of each trial. This removal, as justified in Nozaradan et al. (2011), discards the evoked potential related to the stimuli onset and relies on the fact that the steady-state requires several repetitions or cycles to be elicited (Repp, 2005, Vialatte et al., 2010). In order to enhance the signal-to-noise ratio and attenuate activities that are not phase locked to the auditory and visual stimuli, the EEG epochs for each participant, block, and condition were averaged across trials. We then applied two different frequency analyses: a fast Fourier transform and a periodogram. These transformations yielded a frequency spectrum of the signal's amplitude (µV) or power (μV^2) respectively, ranging from 0 to 500 Hz and showing a frequency resolution of 0.0305 Hz.

The signal obtained from the frequency spectra is assumed to correspond to the EEG activity induced by the physical stimuli and the meter imagery. However, the signal may also include residual background noise because of spontaneous activity. Thus, two different signal-to-noise techniques were applied to each frequency analysis: first, the subtraction method used in Nozaradan et al. (2011) for the amplitude spectrum and, second, a relative measure that compared binary and ternary meter to the beat baseline for the power spectrum (obtained in decibels).

For the amplitude spectrum, we assume, as Nozaradan et al. (2011), that "in the absence of a steady-state evoked potential, the signal amplitude at a given frequency bin should be similar to the signal amplitude of the mean of the surrounding frequency bins". Accordingly, noise was removed by subtracting the averaged amplitude of the two surrounding non-adjacent frequency bins, ranging from -0.15 to -0.09 Hz and from 0.09 to 0.15 Hz, at each frequency bin between 0.5 to 5 Hz.

For the power spectrum, we assume that the control condition (the beat) can work as a baseline, since the subjects were passively listening to the stimuli and that the EEG activity corresponds to the processing of the perceptual beat without any top-down projection of meter. With these two premises in mind, we took the power spectrum of each meter imagery condition for each participant and divided it by their own control condition. Subsequently, we converted these values into dB by taking the $\log_{10}$ of each value and multiplying by 10. Following this method, we do not need to compare the selected frequency bins among conditions and blocks against the control, but instead check whether the obtained values at each frequency of interest differ from zero. This procedure makes the comparison between modalities more reliable, as the effect of meter is relative rather than absolute. We also used the power spectrum in a third analysis in which a generalized linear model (GLM) was applied, but in this case we obtained the values using the signal-to-noise subtraction method from Nozaradan et al. (2011). In this case, however, we converted our units (μV^2) into decibels (dB) by taking the $\log_{10}$ of each value and multiplying by 10 before applying the signal-to-noise subtraction.

6. Statistical analyses

In order to correct for spectral leakage from our target frequencies, the value corresponding to each frequency was calculated by averaging the three frequency bins centered on the target frequencies. That is, we averaged every target frequency bin ±0.0305 Hz. This leakage correction was applied to the amplitude spectrum and the power spectrum used for the GLM.

It was unnecessary to apply this technique to the original power spectrum because after using the baseline correction method, we saw that the induced activity was centered very concisely in a single bin. The mean of all the participants was calculated for each condition and block in each target frequency. This allowed us to clearly detect the peaks of the target frequencies. For the amplitude spectra, the values for each target frequency (0.8, 1.2, 2.4, 4.8 Hz) were separately submitted to a two-way repeated measures ANOVA with the factors Block (auditory and visual) and Condition (beat, binary, ternary). Subsequent *t*-tests were applied when one of the ANOVA factors was significant. For the power spectra, one sample *t*-tests were used to see if the values at each target frequency significantly differed from zero. The significance level was set at $p < 0.05$ for all statistical analyses.

The present approach does not deal with any topological effect, because our hypothesis does not predict any region of interest for bimodal meter induction, although certain correlations could appear (such as occipital areas showing a strong connection to the visual modality of the stimuli). This is the reason why the means from every electrode were averaged across the scalp, thus excluding selection biases.

III. RESULTS

1. Amplitude spectra

In Figure 4, the mean of all participants' amplitudes (the red line) is plotted over each individual's amplitude spectra (the blue lines) at the target frequencies (0.8, 1.2, 2.4, 4.8 Hz). A clear peak appears at the frequency of the stimuli (2.4 Hz) in all the conditions (beat, binary, ternary). The same applies to the peak of the first harmonic (4.8 Hz), present in all three conditions. Interestingly, the peak at 1.2 Hz only appears in the auditory binary condition (A_2), while the peak at 0.8 Hz is found in both auditory and visual ternary conditions (A_3 and V_3). Furthermore, these plots show a larger peak at the beat frequency for all three auditory conditions compared to their visual analogues, whereas the inverse effect occurs at the frequency of the first harmonic, depicting a larger peak for all three visual conditions.

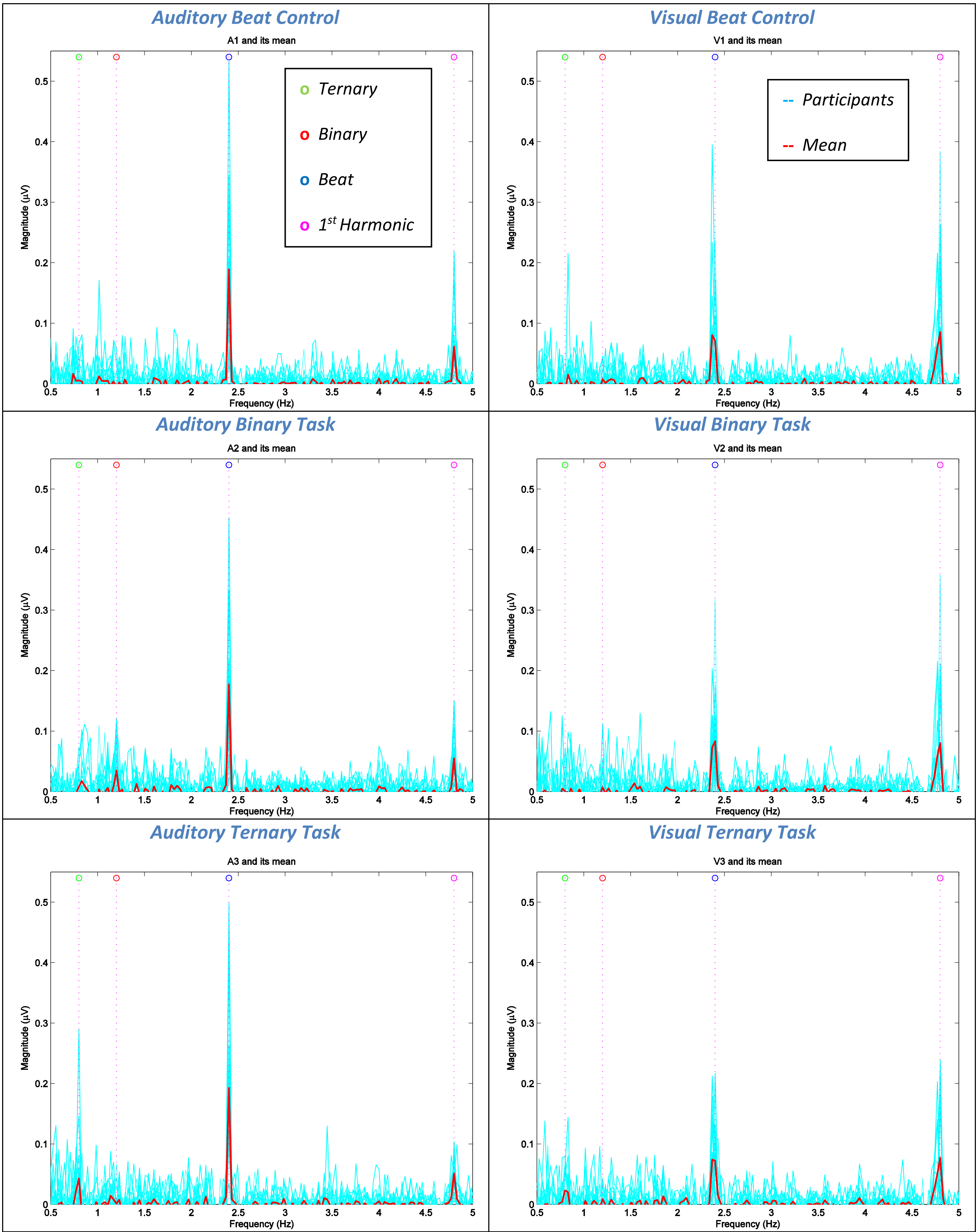


Figure 4. Six amplitude spectra depicting the amplitude (µV) of the averaged EEG signal at each frequency (Hz) between 0.5 and 5 Hz. The auditory (left column) and visual (right column) modalities are split into three conditions: the beat control (first row), the binary meter imagery task (second row) and the ternary meter imagery task (third row). The frequencies of the ternary meter, binary meter, beat and its first harmonic are signaled with dotted lines: 0.8 Hz (green), 1.2 Hz (red), 2.4 Hz (blue) and 4.8 Hz (magenta), respectively.

A two-way repeated measures ANOVA [Block (auditory and visual) x Condition (beat, binary, ternary)] were applied to each frequency of interest (0.8, 1.2, 2.4, 4.8 Hz) separately. For the frequency of the beat (2.4 Hz) and its first harmonic (4.8 Hz), there was a main effect of modality: $F_{(1,15)} = 14.129$, $p = 0.0019$ and $F_{(1,15)} = 10.254$, $p = 0.0059$ respectively. For the ternary subharmonic of the beat ($f/3 = 0.8$ Hz), a main effect of condition was obtained ($F_{(2,30)} = 6.794$, $p=0.0037$). However, for the binary subharmonic of the beat ($f/2 = 1.2$ Hz), no main effect of condition was found ($F_{(2,30)} = 0.774$, $p = 0.47$). Subsequent paired *t*-tests revealed larger amplitude in the auditory modality for all conditions at 2.4 Hz (all *p*-values < 0.011) and larger amplitude in the visual modality for all conditions at 4.8 Hz (all *p*-values < 0.028). There was a significantly larger amplitude in the visual modality in the ternary condition at 0.8 Hz compared to all other conditions except the auditory beat control (p = 0.111, all other *p*-values < 0.031) and in the auditory modality in the ternary condition at 0.8 Hz compared to all other condition except the auditory beat control (p = 0.879, all other *p*-values < 0.027).

These results suggest that there is a similar top-down meter effect for both modalities in the ternary condition enhancing the subharmonic of the beat ($f/3$), the downbeat of a ternary meter occurring at 0.8 Hz. As the paired *t*-tests reported, the amplitude of the enhanced ternary meter in both modalities was significantly greater than in the other modalities except for the auditory control. This finding could be due to the great variance shown by all participants. The same may apply to the peak at 1.2 Hz of the auditory binary condition, which did not turn out significant. The following analyses will carefully check this point by using a different signal-to-noise method.

2. Power spectra

The power spectra for each block and condition were obtained by taking the modulus squared of the amplitudes resulting from the fast Fourier transform. However, here we used the control (beat) condition as a baseline to normalize and convert the amplitudes from the metrical conditions (binary, ternary) into decibels. This procedure consisted of applying the following operation: $10\log_{10}\frac{condition}{control}$. This method gives us the opportunity to see relative distances from zero as differences between conditions with respect to the control. When positive, the values indicate more power for the metrical condition, whereas when negative, they indicate more power for the beat baseline. Figure 5 shows the frequency spectrum with all the electrodes for the auditory binary task and reveals positive activity at 1.2 Hz (see all other conditions in Appendix 1).

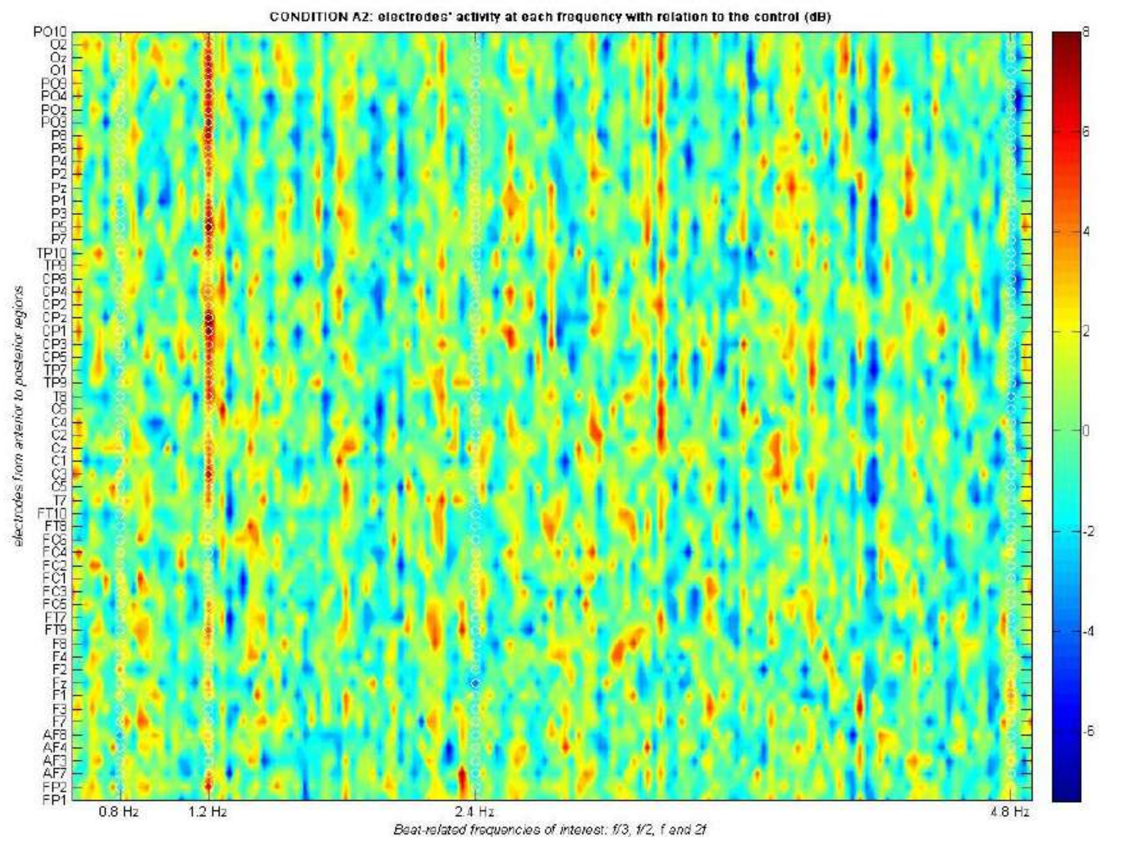


Figure 5. At each electrode, the relative value in dB obtained after subtracting the power spectrum (µV2/Hz) of the control condition from the auditory binary condition. The electrodes are ordered from anterior to posterior, with odd numbers representing the right hemisphere and even ones the left. The frequencies of interest (0.8, 1.2, 2.4, and 4.8 Hz) are signaled at the frequency spectra (from 0.6 to 5 Hz) with white line of dots. The relative power goes from -7 dB (greater in the control condition) to 8 dB (greater in the metrical condition).

To summarize the differences between conditions and blocks, Figure 6 depicts the mean of all the participants contrasting each condition within their modality. The frequencies of interest are marked with a vertical line to make it easier to see the meter-induced peaks at the subharmonics of the beat. Similar to the findings in the amplitude spectra, there are larger peaks at 0.8 Hz for the ternary condition in both modalities, but there is a peak at 1.2 Hz for the binary condition only in the auditory modality. In contrast with our first analyses, no differences arise between modalities at the frequency of the beat (2.4 Hz) in the power spectra. This is due to the use of the beat condition as a baseline. Given that beat is similarly induced in all conditions, dividing the two conditions will result in a relative value close to zero.

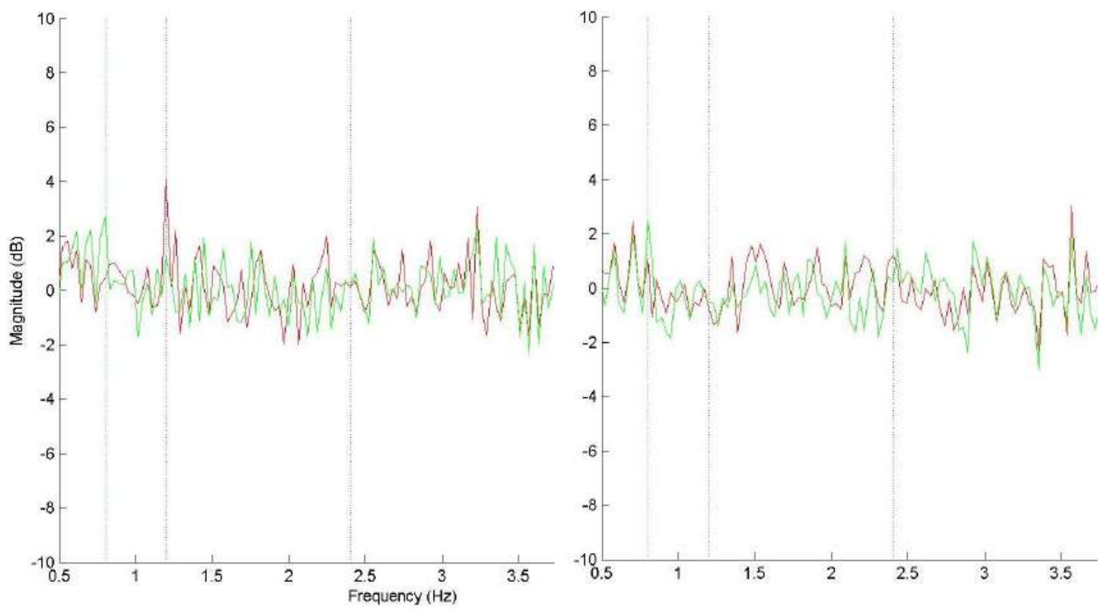


Figure 6. The normalized power spectrum of the ternary and the binary conditions separated into blocks. In both auditory (left) and visual (right) modalities, a peak for the ternary condition (green) appears at 0.8 Hz (green dotted line). A peak is only found in the auditory modality for the binary condition at 1.2 Hz (red dotted line). There is no contrasting peak at 2.4 Hz (blue dotted line) in any modality.

One-sample *t*-tests were applied to the spectrum of each target frequency to examine whether the values significantly differed from zero. At 0.8 Hz, the values of the ternary conditions were significantly different from zero in the auditory ($t_{(15)}$ = 2.778, p = 0.014) and the visual ($t_{(15)}$ = 3.692, p = 0.002) modalities. The same occurred at 1.2 Hz for the values of the binary condition in the auditory modality ($t_{(15)}$ = 3.489, p = 0.003). These findings show that the peaks appearing in both ternary conditions at 0.8 Hz and the auditory binary condition at 1.2 Hz are significantly different from zero and thus larger than the control condition. However, no effects for the visual binary meter were attested, likely due to limitations of our stimuli (see Discussion).

3. Generalized Linear Model

Before applying a generalized linear model (GLM) to the data obtained from the power spectrum using the periodogram function in Matlab, we needed to corroborate the effects reported by the two-way repeated measures ANOVAs run on the amplitude spectra. This was done in order to assure that the effects found after using the subtraction technique employed by Nozaradan et al. (2011) were not dissimilar. All the electrodes were averaged for each block and condition. Figure 7 displays every participant's magnitude at the target frequency for each condition. In line with our previous statistical analyses, one can see the reverse pattern of the values at 2.4 and 4.8 Hz: 2.4 Hz shows larger magnitudes compared to 4.8 Hz for the auditory modality, while this pattern is the opposite for the visual modality (4.8 Hz larger than 2.4 Hz). Furthermore, a small increase of the 0.8 Hz values in the ternary condition also reflects the effect of mentally projecting a ternary structure onto the external stimuli shown by the previous analyses. We then applied a two-way repeated measures ANOVA to each frequency as before. Significant main effects of condition (beat, binary, ternary) were found at 0.8 Hz ($F_{(1,15)}$ = 4.068, p = 0.027), as well as main effects of modality at 2.4 ($F_{(1,15)}$ = 7.496, p = 0.015) and 4.8 Hz ($F_{(1,15)}$ = 8.139, p = 0.012).

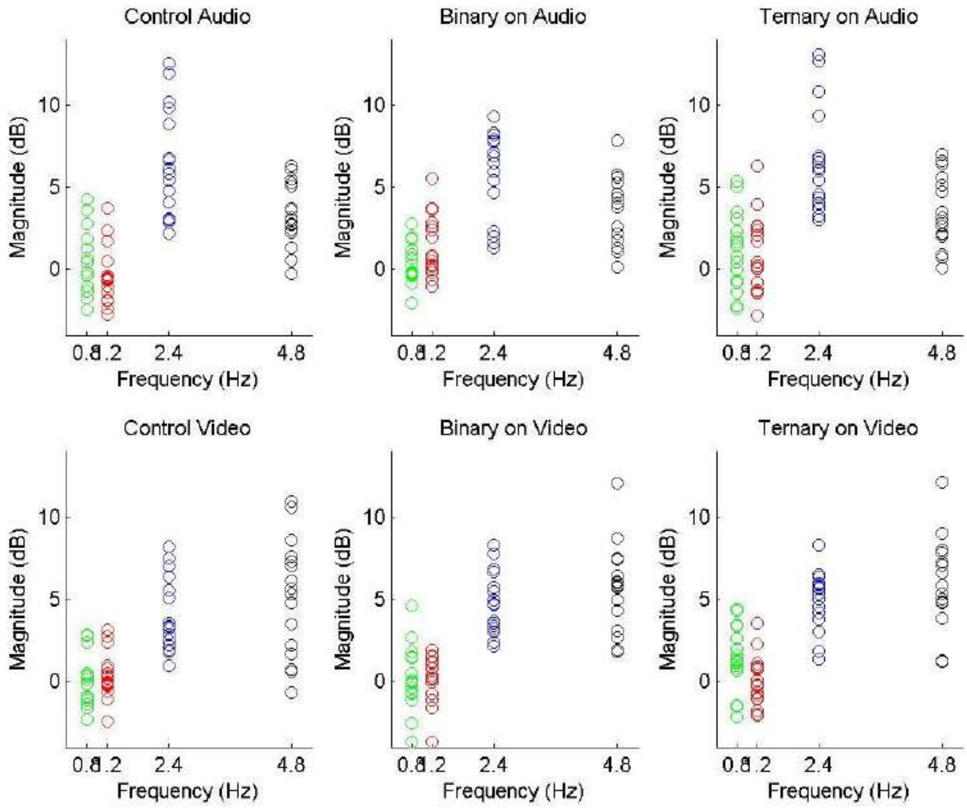


Figure 7. For each condition and block, the relative magnitude of each participant's periodogram at the frequencies 0.8, 1.2, 2.4 and 4.8 Hz. In both auditory (first row) and visual (second row) modalities, the ternary condition shows an increase of each participant's magnitude at 0.8 Hz (green dots). The effect of the binary condition at 1.2 Hz (red dots) only seems to be enhanced in the auditory modality. While the magnitudes at 4.8 Hz (black dots) are smaller than those at 2.4 Hz (blue dots) in the auditory modality, this relation is the opposite in the visual modality.

Having shown that the results found after applying Nozaradan et al.'s subtraction method were the same, we applied a GLM to investigate other factors that could have affected our results. We tested whether these magnitudes also depended on other predictors, making use of the personal data collected through the questionnaires (see Appendix 3). We selected the following predictors: binary meter imagery, ternary meter imagery, visual block, years of music training, dance training, and the interaction between visual modality and dance (see Table 1).

For our hypothesis, it is relevant to look at the dance factor because this activity involves extracting the beat from the visual domain when group synchronization is required.

	0.8Hz		**1.2Hz**		**2.4Hz**		**4.8Hz**	
	beta	***p-value***	***beta***	***p-value***	***beta***	***p-value***	***Beta***	***p-value***
	0,3398	0,7088	-1,6025	0,0609	2,9468	0,0135	-0,3623	0,7551
Binary meter	0,0797	0,8664	0,8001	0,0724	0,0469	0,9389	0,3544	0,5588
Ternary meter	**1,0436**	**0,0299**	0,4405	0,3196	0,4715	0,4412	0,4880	0,4212
Visual modality	-0,2249	0,6633	-0,4482	0,3521	-0,7185	0,2818	**1,8343**	**0,0064**
Years of music	0,0045	0,9232	**0,1067**	**0,0155**	**0,1399**	**0,0217**	**0,2004**	**0,0011**
Dance training	-0,2996	0,5875	0,1149	0,8231	**1,9961**	**0,0060**	0,3267	0,6432
*Visual*Dance*	0,2012	0,7966	-0,1061	0,8838	**-2,4254**	**0,0176**	0,6587	0,5092

Table 1. GLM applied to each target frequency (0.8, 1.2, 2.4, and 4.8 Hz). The betas and p-values are shown at each target frequency for the following predictors: binary meter task, ternary meter task, visual modality, years of music training, dance training, and the interaction between the visual modality and dance training. Significant values are shown in bold.

The results of the GLM are displayed in Table 1. At 0.8 Hz, the GLM reveals ternary meter as a significant predictor ($\beta = 1.044$, $p = 0.030$). At 1.2 Hz, years of music training is a significant predictor ($\beta = 0.107$, $p = 0.016$). At 2.4 Hz, years of music training ($\beta = 0.14$, $p = 0.022$), dance training ($\beta = 1.996$, $p = 0.006$), and the visual modality x dance training interaction ($\beta = -2.425$, $p = 0.018$) are significant predictors. At 4.8 Hz, visual modality ($\beta = 1.834$, $p = 0.006$) and years of music training ($\beta = 0.200$, $p = 0.001$) are shown to be significant.

Because we found a significant interaction of visual modality and dance training at 2.4 Hz, we applied a GLM to the dancers and the non-dancers separately (see Table 2). This table shows the visual modality to be significant at 2.4 Hz only for dancers ($\beta = -1.687$, $p = 0.010$). Since beta is negative, it implies that the magnitude at 2.4 Hz is reduced for the visual modality. In contrast, at the same frequency, non-dancers have years of music training as a significant predictor ($\beta = 0.238$, $p = 0.003$): the more time spent training, the larger the magnitude. Though not pertinent to the 2.4 Hz interaction that was found, we looked at the other target frequencies as well. At the first harmonic (4.8 Hz), both populations (dancers and non-dancers) seem to be significantly affected by the visual modality ($\beta = 0.686$, $p = 0.001$ and $\beta = 1.404$, $p = 0.000$ respectively), now positively, thus showing increased magnitude at 4.8 Hz. Curiously, the dancer group does not show ternary condition as significant at 0.8 Hz ($\beta = 0.083$, $p = 0.442$), whereas the ternary condition is significant for the non-dancers ($\beta = 0.760$, $p = 0.010$).

	0.8Hz		1.2Hz		2.4Hz		4.8Hz	
	beta	*p-value*	*beta*	*p-value*	*beta*	*p-value*	*Beta*	*p-value*
DANCERS	-0,0612	0,7056	-0,0953	0,3922	2,5821	0,0225	0,1061	0,7613
Binary meter	0,1068	0,3219	0,0686	0,3529	-0,0946	0,8958	-0,1086	0,6386
Ternary meter	0,0825	0,4427	0,0104	0,8869	0,0660	0,9272	-0,1356	0,5579
Visual modality	-0,0705	0,4219	-0,1100	0,0725	**-1,5873**	**0,0101**	**0,6856**	**0,0008**
Years of music	0,0034	0,6812	0,0089	0,1175	-0,0191	0,7291	0,0110	0,5330
NON-DANCERS	0,0334	0,9632	-0,4328	0,1496	-2,2204	0,1087	-1,2523	0,2474
Binary	0,0986	0,7305	0,1636	0,1684	-0,2665	0,6226	-0,1862	0,6621
Ternary	**0,7599**	**0,0103**	0,0447	0,7042	-0,2509	0,6431	-0,3691	0,3878
Visual	-0,1234	0,5979	-0,0602	0,5319	0,2153	0,6262	**1,4036**	**0,0002**
MusicYears	-0,0036	0,9299	0,0253	0,1332	**0,2381**	**0,0030**	0,1052	0,0855

Table 2. GLM separately applied to dancers and non-dancers at each target frequency (0.8, 1.2, 2.4, and 4.8 Hz). The betas and p-values are shown at each target frequency for the following predictors: binary meter task, ternary meter task, visual modality, and years of studying music. Significant value are shown in bold.

IV. DISCUSSION AND FURTHER RESEARCH

In the present work, we studied meter induction in the auditory and visual modalities by means of electroencephalography (EEG). Our aim was to see whether the hierarchical organization of beats (meter) is also present in the visual domain. We analyzed the data via two different signal-to-noise procedures applied to two types of frequency analyses: amplitude and power spectra. These different approaches allowed us to reliably compare the conditions between and across modalities.

Our work is among the first studies on sensorimotor synchronization (SMS) that compares meter induction between two different modalities, audition and vision, without requiring overt movement coordinated with an external rhythm. As far as we know, SMS cross-modal studies have mainly explored the degree to which one can reliably extract the pulse from visual rhythms (Repp, 2005, Repp and Su, 2013, Patel et al. 2005). One such cross-modal study (Phillips-Silver and Trainor, 2005) investigated metrical preferences in infancy by bouncing the babies at a binary or ternary pulse, thus priming their vestibular system. However, in contrast to these studies, our work does not depend on motor acts to induce meter in the visual domain.

1. Beat resonance

Our results corroborate, as Nozaradan et al. (2011) suggested, that EEG recordings are a perfect tool to capture periodic responses in the brain, which, in our case, were steady-state evoked potentials elicited by the entrainment of neuronal populations to the frequencies of the beat and the meter. The interesting point is the emergence or enhancement of periodic responses at the subharmonics ($f/2$ and $f/3$) of the beat (f = 2.4 Hz) in the binary and the ternary auditory conditions. Hence, the EEG recordings seem to reflect the voluntary metrical interpretation of the beat, despite the fact that the stimuli presented in the different conditions (binary or ternary) were identical. Furthermore, since we obtained similar results for the visual and the auditory condition, we can claim that neural entrainment not only occurs to the beat in the visual modality, but also to the subharmonic $f/3$ in the visual ternary condition, which supports top-down effects of meter induction in the visual modality.

The present results fit well into the framework of the dynamic attending theory (Jones, 1976; Jones and Boltz, 1989; Large and Jones, 1999), which considers that participants' attention entrains and fluctuates following the periodic structure of the stimuli, thus periodically modulating their expectancy as a function of time. According to the dynamic attending theory, "tone sequences presented at a regular rhythm entrain attentional oscillations and thereby facilitate the processing of sounds presented in phase with this rhythm" (Bauer et al., 2015). Furthermore, our beat- and meter-induced periodicities also fit the resonance theory of beat and meter (Large, 2008; Large and Snyder, 2009), although it seems to apply beyond the auditory modality. Meter, as an endogenous representation of the beat, may therefore be a higher-order resonance product which can be modulated voluntarily. This modulation may be achieved by shifting our focus of attention to a chosen subharmonic of the beat, thus reinforcing the neural activity tuned to the selected periodicity.

2. Meter induction

In our analyses, we did not observe any effect of the binary meter in the visual modality. We did, however, find an effect of the ternary meter in this modality. This can be interpreted as a cue that meter induction applies to the visual domain independently of a conversion between the visual and the auditory modalities. Our argument assumes that equal effects for binary and ternary conditions should be expected regardless of whether visual information was converted into auditory information[4] or not. If we had found meter-induced effects in both visual conditions, we would not have been able to distinguish whether it was due to an automatic conversion of the visual beat into auditory information or an actual cross-modal effect of meter induction. Had we not found any meter-elicited SS-EP for the visual ternary condition, we would have logically concluded that neither visual meter induction nor visual-to-auditory information conversion were in place. Nevertheless, our results suggest that visual cues could have interacted with a more complex integrative process. More relevant to our work, the results from the visual modality suggest that meter induction applies to the induced beats, regardless of its perceptual nature.

The effects of imposing a metrical structure on the periodic stimuli show that humans have the ability to select and enhance certain beat-related harmonics through a dynamic top-down biasing processing. As was reported in the questionnaires, participants used not only counting to feel the meter, but other mechanisms as well, such as imagining variations of intensity (luminance, volume), pitch or tone variation (change in frequency), and even spatial distortion of the flashes following the meter. These enhanced additional periodicities seem to be selected among preferred integer ratios of the beat, such as 3:1 and 2:1. Keller (2012) discusses that the predisposition for a metrical pattern is cultural and learned by experience and that other musical idioms (e.g. Balkan rhythms, Indian ragas, African drumming) can make use of other subharmonics and non-integer ratios leading to complex metrical patterns, such as 5:1, 7:1, 11:1, which are very atypical in Western music.

3. Modality effects

Our analyses revealed the emergence of a periodic entrainment at the frequency of the first harmonic of the beat, at $2f$ (4.8 Hz). This finding illustrates a natural preference for integer harmonics that tend to emerge involuntarily when a periodic stimulus is presented. In fact, it may relate to our predisposition to subdivide the beat into integer harmonics, like duplets (1:2) or triplets (1:3) of eighth notes.

In the findings of our amplitude spectrum analyses, the amplitude of the first harmonic correlated with the modality of the stimulus, thus appearing more strongly (with a significantly higher peak) for the visual than the auditory modality. In contrast, this enhanced amplitude at

[4] Some authors (Guttman et al., 2005) have proposed that visual cues alternating in duration may be translated into auditory cues in an obligatory and automatic way. McCauley and Henry (2010), however, reported some findings against the strength of this automaticity.

4.8 Hz in the visual conditions was inversely found at the beat frequency (2.4 Hz) for the auditory conditions. Although the reasons for these results in the magnitude spectrum are still unclear, one could hypothesize that this opposite effect could be due to the comfortable frequencies at what vision and audition tend to work. Another possibility could be that it is more difficult to process our visual stimuli and therefore keeping the beat may lead to a need to compensate by reinforcing the first harmonic.

Contrasting with the previous results, the power spectrum analyses did not show any modality-dependence on either the beat or the first harmonic. We should keep in mind, however, that a signal-to-noise procedure dividing the control values over each condition as a baseline was applied to the values of the power spectrum. This procedure makes them relative and comparable but, at the same time, devoid of their absolute dimension. In other words, the absolute power for each condition and frequency was lost after normalization. This means that the beat-enhanced frequency power, and even its first harmonic power, tends to zero because there are no differences between the control condition and the metrical conditions related to the perception of the external stimuli.

In the generalized linear model, our findings reveal an interaction between dance training and visual conditions. Dancers indeed showed a significantly smaller effect (expressed by less dB) for visual conditions than non-dancers. Curiously, when the GLM was applied separately to these two populations, the amount of training in years as a significant predictor for beat synchronization was only maintained for non-dancers. The same occurred to the effect of meter at 0.8 Hz for the ternary condition, which was lost for dancers. What is more relevant here is that dance could have an influence on visuomotor engaging (i.e. the coupling of perceived visual movements onto motor acts), because dance embodies several metrical levels of the beat within distinct movement dimensions (Toiviainen et al., 2010), such as limb, trunk and head periodic movements (Leman and Naveda, 2010). Hence, one conclusion could be that visual beat induction (but not meter) would be facilitated by dance, requiring less effort (i.e. less neural activity tuned to the beat frequency). However, since the ternary condition for dancers does not significantly predict the metrical effect enhancement at 0.8 Hz (something that dance training and visual domain succeed to significantly predict), our idea that meter is cross-modal remains an open question for this population.

4. Study limitations

There are some possible explanations of why we did not obtain a metrical enhancement at 1.2 Hz in the visual modality. The first cause could have been the kind of stimuli presented to the participants. As will be discussed below, other kinds of visual stimuli (biological and spatial) might help to trigger metrical enhancement at different frequencies. However, it is important to remember that the stimuli we used worked very well for the ternary condition. Another possibility is the time-specificity of each modality, that is, the preferred tempo for each modality is not the same (Vialatte, Maurice, Tanaka et al., 2010). We are not claiming that time does not relate more accurately to the auditory than to the visual modality (see the modality-appropriateness hypothesis in Welch, 1999). However, our results support the idea that meter could be a mechanism working on attended frequency subharmonics that aids the processing of periodic stimuli regardless the modality used to perceive them. This supports the idea of meter as a top-down chunking mechanism that can be triggered in perception and production.

Visuomotor synchronization is a current topic of research because of the greater variability in the results of beat tapping with visual rhythms, compared to auditory or tactile rhythms. Hove, Spivey and Krumhansl (2010) justify the difficulty of visual synchronization claiming that it could be due to a lesser degree of environmental visual rhythms, a tighter auditory-motor connected physiology[5], or the inherent specialization of each system for the processing of certain types of information: temporal versus spatial. In contrast to the study by Nozaradan et al. (2011), our study did not require beat induction per se, because both auditory and visual stimuli were presented in a strictly periodic pattern, without being modulated in amplitude (volume/luminance) at a faster frequency. Since we avoided beat induction from visual rhythms to facilitate the abstraction of the beat and thus focus on the metrical effects, we are omitting half of the actual process. Our study, therefore, is limited to demonstrate visual meter induction only for visual periodic beats, which are not extracted from more complex, natural rhythms.

Our visual stimuli were colored flickering flashes that suddenly appeared and vanished to disappear just before it appeared again. This kind of stimulus has been extensively used to create steady-state evoked potentials (Vialatte et al., 2010; Krause, Pollock and Schnitzler, 2010). However, some studies have claimed to get better results from the visual domain using geometrical and spatial features (Hove, Spivey and Krumhansl, 2010), as well as biological motion and natural speed effects (Su, 2014a, 2014b; Hove, Iversen et al. 2013). These stimuli have been found to improve beat synchronization and influence rhythm perception. In fact,

---

[5] An auditory advantage for SMS was proposed (Grahn et al., 2011; Zatorre et al., 2007; Merchant et al., 2015) based on the stronger coupling between the motor and auditory areas. Besides, even a hypothesis on "action simulation for auditory prediction" (Patel and Iversen, 2014), based on simulations of periodic movements, also considers that beat and meter arise because of the "temporally-precise communication between auditory regions and motor planning regions of the cortex".

Gan et al. (2015) obtained even slightly better synchronization results for the IOI from 500 to 900ms by using a depicted bouncing ball with realistic motion trajectory as opposed to a conventional auditory metronome. Moreover, their behavioral findings also suggest that there is no modality bias of beat synchronization. In the same direction, Trost et al. (2014) showed cross-modal attention effects due to musical meter in the caudate nucleus, and Hove, Fairhurst et al. (2013) found that the basal ganglia activity was more associated to SMS stability than to modality features. Thus, the biological substrates of beat synchronization should be considered more independently from the input modality. It would be interesting to explore the effects of using these kinds of stimuli in a task such as the one used in our study.

5. Further issues

One issue that needs discussion is the extent to which the present findings can be generalized across different populations. In our study, we tested highly proficient musicians. Kung et al. (2011) reported that musicians, compared to non-musicians, located more precisely on time clicks that coincided with a strong beat at the end of a rhythmic group, and even better when visual cues were removed. This indicates that "musical expertise not only enhances attention to metrical accents but also heightens sensitivity to perceptual grouping." Also relevant for us is what Repp (2010) reported when exploring illusory phenomenal accents: musicians showed "an increase of sensitivity to physical changes in main beat positions, likely to be due to enhanced attention." This finding would give more support to the dynamic attending theory mentioned before (Jones and Boltz, 1989; Large and Jones, 1999), in which meter may allow for a cyclical fluctuation of attention over isochronous events, yielding expectancies and predictions of incoming beats.

Future research using a mismatch negativity (MMN) design could shed light on visual meter induction, similar to what Honing et al. (2014) propose for beat induction. Participants could be asked to mentally project a metrical structure on sequences of flashes. Some flashes coinciding with strong and weak positions would be omitted. Then, a comparison of the obtained effects would support or reject our findings on cross-modal meter induction. If different electrophysiological responses were attested for flashes occurring at the strong position, it would provide further evidence that meter applies to the visual modality. Moreover, new visual stimuli should be created to facilitate visual beat perception and therefore compare the magnitudes of effort-equilibrated tasks. The use of naturally-driven movements seems to be promising for this purpose.

V. CONCLUSION

Our study supports that meter induction is an endogenously driven percept, which does not require external perceptual cues to be elicited. As a top-down mechanism, it mediates the way that a beat is elicited by a stimulus regardless of the modality used to perceive it, in our case audition and vision. Since we found similar effects on the auditory and the visual domains for the ternary imagery of the meter, but not for the binary projected meter, we may conclude that we are not dealing with a pure conversion of visual features into auditory ones. Our results suggest that meter should be conceptualized as a cross-modal, attentional mechanism available in both domains to deal with and focus on certain relevant periodicities, such as the subharmonics of a given beat. Future research on attention and visuomotor integration will reveal how a cross-modal meter could be implemented in neural terms, which would also be relevant for understanding group synchronization or dance.

VI. 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*.

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),362-6.

Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body Movement Selectively Shapes the Neural Representation of Musical Rhythms. *Psychological science*, 0956797614551161.

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in systems neuroscience*, *7*.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see auditory encoding of visual temporal sequences. *Psychological Science*, *16*(3), 228-235.

Gan, L., Huang, Y., Zhou, L., Qian, C., & Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. *Scientific reports*, *5*.

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *Neuroimage*, *54*(2), 1231-1243.

Honing, H., Bouwer, F. L., & Háden, G. P. (2014). Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception. In *Neurobiology of interval timing* (pp. 305-323). Springer New York.

Hove, M. J., & Keller, P. E. (2010). Spatiotemporal relations and movement trajectories in visuomotor synchronization.

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1525.

Hove, M. J., Iversen, J. R., Zhang, A., & Repp, B. H. (2013). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome.*Psychological research*, *77*(4), 388-398.

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *Neuroimage*, *67*, 313-321.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it?. *Cognition*, *100*(1), 33-72.

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., & Boltz, M. (1989). Dynamic attending and responses to time.*Psychological rev.*, *96*(3), 459.

Keller, P. E. (2012). Rhythm and Time in Music Epitomize the Temporal Dynamics of Human Communicative Behavior: The Broad Implications of London's Trinity.

Krause, V., Schnitzler, A., & Pollok, B. (2010). Functional network interactions during sensorimotor synchronization in musicians and non-musicians.*Neuroimage*, *52*(1), 245-251

Kung, S. J., Tzeng, O. J., Hung, D. L., & Wu, D. H. (2011). Dynamic allocation of attention to metrical and grouping accents in rhythmic sequences.*Experimental brain research*, *210*(2), 269-282.

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., & Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychological review*, *106*(1), 119.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance.*Annals of the New York Academy of Sciences*, *1169*(1), 46-57.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception*, 229-252.

Naveda, L., & Leman, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (TGA).

Large, E. W. (2008). Resonating to Musical Rhythm: Theory and Experiment. In: S. Grondin (Ed.), Psychology of time (pp. 189- 231). Bingley, UK: Emerald Group Publishing Limited.

Lerdahl, F., & Krumhansl, C. L. (2007). MODELING TONAL TENSION. *Music Perception*, *24*(4), 329.

Leman, M., & Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston, 71-91.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.

McAuley, J. D., & Henry, M. J. (2010). Modality effects in rhythm processing: auditory encoding of visual rhythms is neither obligatory nor automatic.*Attention, Perception, & Psychophysics*, *72*(5), 1377-1389.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *370*(1664), 20140093.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234-10240.

Nozaradan, S. (2013). Exploring the neural entrainment to musical rhythms and meter: A steady-state evoked potential approach. *Unpublished doctoral thesis, Université Catholique de Louvain, Belgium, and Université de Montréal, Quebec, Canada*.

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1658), 20130393.

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. Cerebral Cortex, 25(3), 736-747.

Overy, K. (2012). Musical Rhythm for Linguists: A Response to Justin London.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in systems neuroscience*, *8*.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*(2), 226-238.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of motor behavior*, *35*(4), 355-370.

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic bulletin & review*, *12*(6), 969-992.

Repp, B. H. (2010). Do metrical accents create illusory phenomenal accents?.*Attention, Perception, & Psychophysics*, *72*(5), 1390-1403.

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

Su, Y. H. (2014a). Audiovisual beat induction in complex auditory rhythms: point-light figure movement as an effective visual beat. *Acta psychologica*,*151*, 40-50.

Su, Y. H. (2014b). Peak velocity as a cue in audiovisual synchrony perception of rhythmic stimuli. *Cognition*, *131*(3), 330-344.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement.

Trost, W., Frühholz, S., Schön, D., Labbé, C., Pichon, S., Grandjean, D., & Vuilleumier, P. (2014). Getting the beat: Entrainment of brain activity by musical rhythm and pleasantness. *NeuroImage*, *103*, 55-64.

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2009). Steady state visual evoked potentials in the delta range (0.5-5 Hz). In *Advances in Neuro-Information Processing* (pp. 400-407). Springer Berlin Heidelberg.

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology*, *90*(4), 418-438

Vialatte, F. B., Maurice, M., Tanaka, T., Yamaguchi, Y., & Cichocki, A. (2010). Analyzing steady state visual evoked potentials using blind source separation. In *Proceedings of the Second APSIPA Annual Summit and Conference, Singapore: APSIPA* (pp. 578-582).

Welch, R.B. (1999). Meaning, attention, and the ‘‘unity assumption’’ in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Musseler (Eds.), Cognitive contributions to the perception of spatial and temporal events (pp.371–387). Amsterdam: Elsevier.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547-558.

VII. APPENDIX

1. Amplitude spectra paired *t*-tests

Paired-sample *t*-tests of the ternary subharmonic (1.2 Hz) comparing conditions and blocks: A1, A2, A3, V1, V2, V3

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Condidence Interval of the Difference				
					Lower	Upper			
Par 1	F08_A1 - F08_A2	418,44574	703,88969	175,97242	43,36940	793,52208	2,378	15	,031
Par 2	F08_A1 - F08_A3	-384,85333	910,07984	227,51996	-869,80065	100,09398	-1,692	15	,111
Par 3	F08_A1 - F08_V1	407,29992	816,24260	204,06065	-27,64506	842,24490	1,996	15	,064
Par 4	F08_A1 - F08_V2	333,45392	930,57012	232,64253	-162,41189	829,31974	1,433	15	,172
Par 5	F08_A1 - F08_V3	-25,69341	661,25484	165,31371	-378,05124	326,66443	-,155	15	,879
Par 6	F08_A2 - F08_A3	-803,29908	1312,66278	328,16570	-1502,76770	-103,83045	-2,448	15	,027
Par 7	F08_A2 - F08_V1	-11,14582	613,11016	153,27754	-337,84916	315,55752	-,073	15	,943
Par 8	F08_A2 - F08_V2	-84,99182	791,44012	197,86003	-506,72049	336,73685	-,430	15	,674
Par 9	F08_A2 - F08_V3	-444,13915	621,49620	155,37405	-775,31110	-112,96720	-2,859	15	,012
Par 10	F08_A3 - F08_V1	792,15325	1260,28262	315,07066	120,59605	1463,71046	2,514	15	,024
Par 11	F08_A3 - F08_V2	718,30725	1122,73633	280,68408	120,04330	1316,57121	2,559	15	,022
Par 12	F08_A3 - F08_V3	359,15992	942,96601	235,74150	-143,31119	861,63104	1,524	15	,148
Par 13	F08_V1 - F08_V2	-73,84600	765,83961	191,45990	-481,93312	334,24112	-,386	15	,705
Par 14	F08_V1 - F08_V3	-432,99333	698,86496	174,71624	-805,39218	-60,59448	-2,478	15	,026

Paired-sample *t*-tests of the binary subharmonic (1.2 Hz) comparing conditions and blocks: A1, A2, A3, V1, V2, V3

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the difference Lower	Upper			
Par 1 F12_A1 - F12_A2	-28,47419	926,98836	231,74709	-522,43142	465,48304	-,123	15	,904
Par 2 F12_A1 - F12_A3	-180,96491	1003,82818	250,95704	-715,86719	353,93737	-,721	15	,482
Par 3 F12_A1 - F12_V1	159,63454	701,11083	175,27771	-213,96105	533,23014	,911	15	,377
Par 4 F12_A1 - F12_V2	184,04661	749,60550	187,40138	-215,38996	583,48319	,982	15	,342
Par 5 F12_A1 - F12_V3	34,45226	880,39727	220,09932	-434,67833	503,58285	,157	15	,878
Par 6 F12_A2 - F12_A3	-152,49072	773,75439	193,43860	-564,79533	259,81389	-,788	15	,443
Par 7 F12_A2 - F12_V1	188,10873	765,45740	191,36435	-219,77472	595,99219	,983	15	,341
Par 8 F12_A2 - F12_V2	212,52080	683,40479	170,85120	-151,63991	576,68151	1,244	15	,233
Par 9 F12_A2 - F12_V3	62,92645	1018,64313	254,66078	-479,87016	605,72306	,247	15	,808
Par 10 F12_A3 - F12_V1	340,59945	698,18395	174,54599	-31,43651	712,63542	1,951	15	,070
Par 11 F12_A3 - F12_V2	365,01152	706,31321	176,57830	-11,35622	741,37926	2,067	15	,056
Par 12 F12_A3 - F12_V3	215,41717	914,06148	228,51537	-271,65181	702,48616	,943	15	,361
Par 13 F12_V1 - F12_V2	24,41207	382,96853	95,74213	-179,65746	228,48159	,255	15	,802
Par 14 F12_V1 - F12_V3	-125,18228	494,35369	123,58842	-388,60477	138,24021	-1,013	15	,327
Par 15 F12_V2 - F12_V3	-149,59435	779,28192	194,82048	-564,84437	265,65567	-,768	15	,454

Paired-sample *t*-tests of the beat frequency (2.4 Hz) comparing conditions and blocks: A1, A2, A3, V1, V2, V3

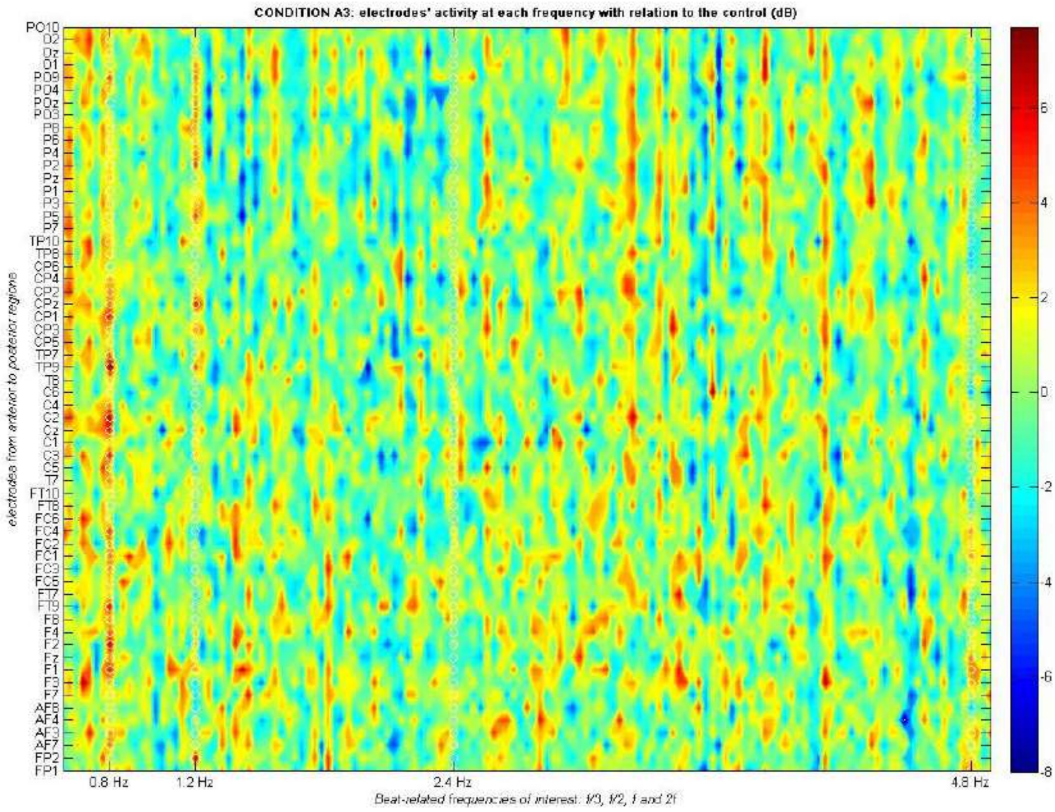
	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	Upper			
Par 1 F24_A1 - F24_A2	14,02609	448,87659	112,21915	-225,16336	253,21554	,125	15	,902
Par 2 F24_A1 - F24_A3	24,49100	607,60259	151,90065	-299,27756	348,25956	,161	15	,874
Par 3 F24_A1 - F24_V1	-865,36805	1091,62783	272,90696	-1447,05546	-283,68064	-3,171	15	,006
Par 4 F24_A1 - F24_V2	-664,14241	848,71924	212,17981	-1116,39296	-211,89185	-3,130	15	,007
Par 5 F24_A1 - F24_V3	-636,10059	875,03324	218,75831	-1102,37289	-169,82829	-2,908	15	,011
Par 6 F24_A2 - F24_A3	10,46491	466,16043	116,54011	-237,93444	258,86427	,090	15	,930
Par 7 F24_A2 - F24_V1	-879,39414	1045,38656	261,34664	-1436,44131	-322,34696	-3,365	15	,004
Par 8 F24_A2 - F24_V2	-678,16849	830,88597	207,72149	-1120,91638	-235,42061	-3,265	15	,005
Par 9 F24_A2 - F24_V3	-650,12668	813,09157	203,27289	-1083,39259	-216,86076	-3,198	15	,006
Par 10 F24_A3 - F24_V1	-889,85905	1116,90222	279,22555	-1485,01423	-294,70387	-3,187	15	,006
Par 11 F24_A3 - F24_V2	-688,63341	948,54234	237,13558	-1194,07594	-183,19087	-2,904	15	,011
Par 12 F24_A3 - F24_V3	-660,59159	723,90836	180,97709	-1046,33512	-274,84805	-3,650	15	,002
Par 13 F24_V1 - F24_V2	201,22564	735,58180	183,89545	-190,73823	593,18952	1,094	15	,291
Par 14 F24_V1 - F24_V3	229,26746	801,33103	200,33276	-197,73171	656,26662	1,144	15	,270
Par 15 F24_V2 - F24_V3	28,04182	565,24724	141,31181	-273,15718	329,24081	,198	15	,845

Paired-sample *t*-tests of the first harmonic (4.8 Hz) comparing conditions and blocks: A1, A2, A3, V1, V2, V3

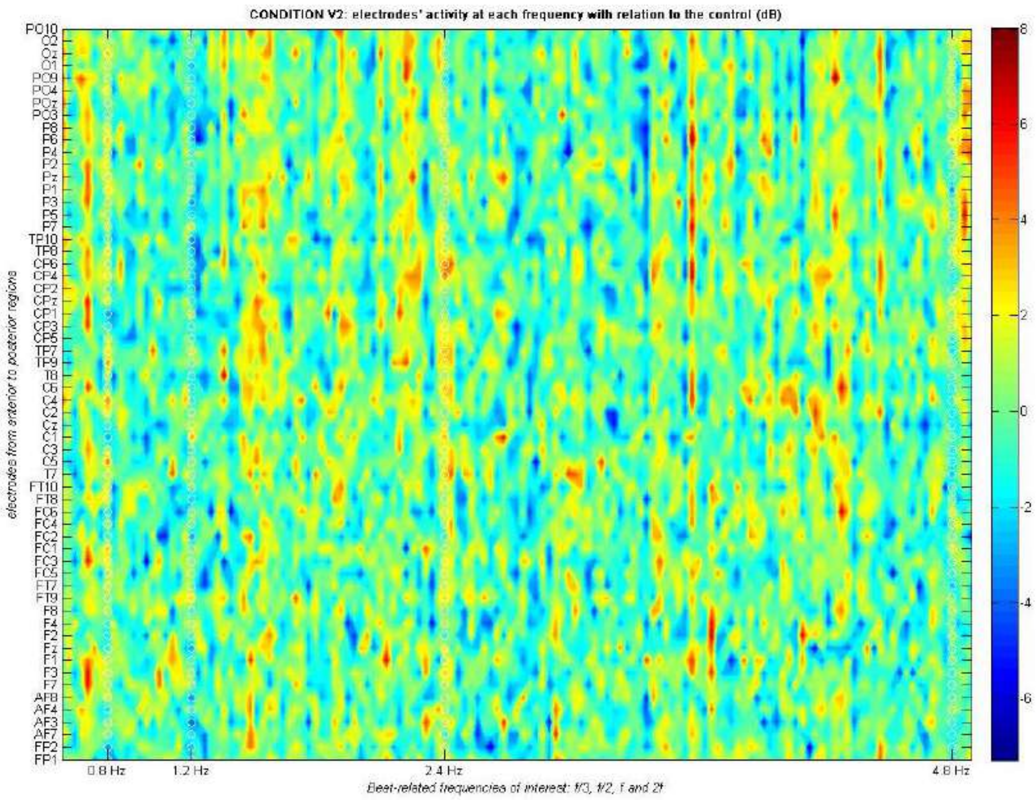
	Paired Samples					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Par 1 F48_A1 - F48_A2	100,46921	385,60915	96,40229	-105,00740	305,94582	1,042	15	,314
Par 2 F48_A1 - F48_A3	-41,59644	464,11003	116,02751	-288,90322	205,71033	-,359	15	,725
Par 3 F48_A1 - F48_V1	-909,39174	1258,84707	314,71177	-1580,18399	-238,59948	-2,890	15	,011
Par 4 F48_A1 - F48_V2	-870,57606	1081,24872	270,31218	-1446,73283	-294,41928	-3,221	15	,006
Par 5 F48_A1 - F48_V3	-823,09035	887,83628	221,95907	-1296,18491	-349,99579	-3,708	15	,002
Par 6 F48_A2 - F48_A3	-142,06566	398,11565	99,52891	-354,20651	70,07520	-1,427	15	,174
Par 7 F48_A2 - F48_V1	-1009,86095	1291,16008	322,79002	-1697,87159	-321,85031	-3,129	15	,007
Par 8 F48_A2 - F48_V2	-971,04527	1108,81374	277,20343	-1561,89040	-380,20014	-3,503	15	,003
Par 9 F48_A2 - F48_V3	-923,55956	1019,90504	254,97626	-1467,02859	-380,09053	-3,622	15	,003
Par 10 F48_A3 - F48_V1	-867,79529	1426,14085	356,53521	-1627,73211	-107,85847	-2,434	15	,028
Par 11 F48_A3 - F48_V2	-828,97961	1277,79096	319,44774	-1509,86635	-148,09288	-2,595	15	,020
Par 12 F48_A3 - F48_V3	-781,49390	1147,24535	286,81134	-1392,81780	-170,17001	-2,725	15	,016
Par 13 F48_V1 - F48_V2	38,81568	349,68281	87,42070	-147,51714	225,14850	,444	15	,663
Par 14 F48_V1 - F48_V3	86,30139	623,70797	155,92699	-246,04913	418,65191	,553	15	,588
Par 15 F48_V2 - F48_V3	47,48571	609,08682	152,27171	-277,07375	372,04517	,312	15	,759

2. Power spectra plots

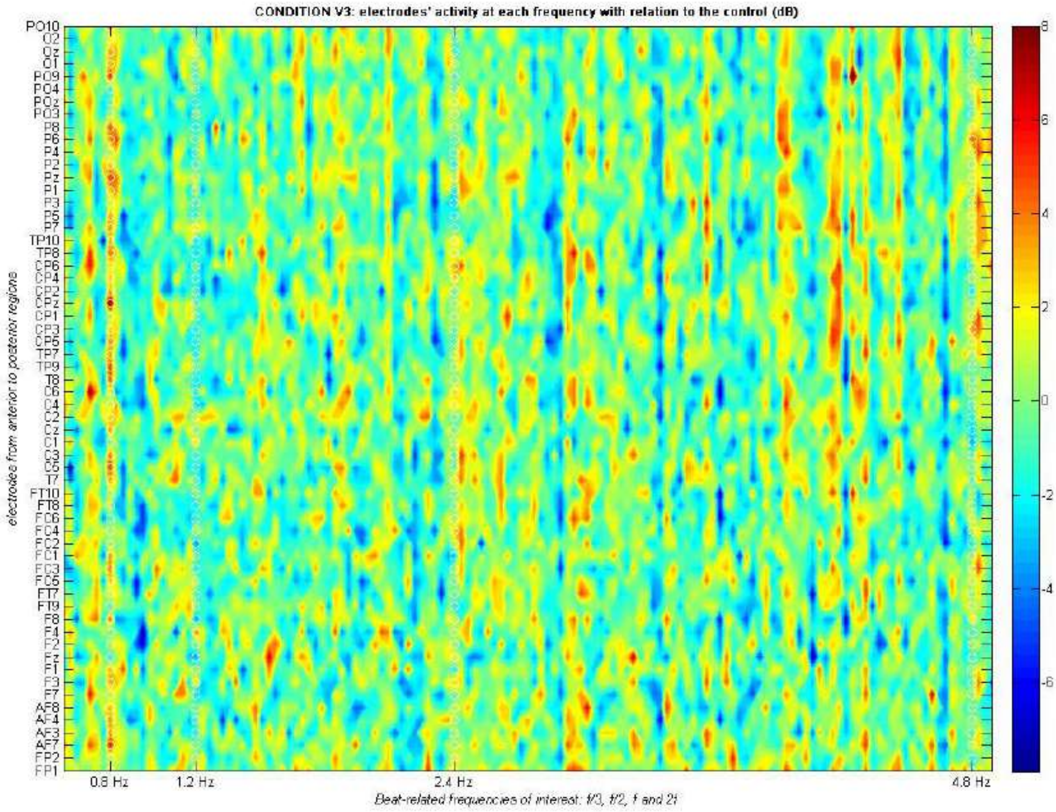
Auditory ternary meter task, each electrode magnitude (dB) over the frequency spectrum



Visual binary meter task, each electrode magnitude (dB) over the frequency-spectrum



Visual ternary meter task, each electrode magnitude (dB) over the frequency-spectrum



3. Questionnaire and evaluation models

Qüestionari per a l'Experiment Musical Rítmic

Cuestionario para el Experimento Musical Rítmico

Sigles del nom i cognoms.

Siglas del nombre y apellidos

Edat

Edad

Sexe

Sexo

Correu electrònic

Correo electrónico

Mà de preferència

Mano de preferencia

Llengua materna

Lengua materna

Quines llengües parleu i domineu?

Qué lenguas habla y domina?

Teniu algun dèficit visual o auditiu? (Ex: otitis, daltonisme, sordesa...)

Tiene algún déficit visual o auditivo? (Ej: otitis, daltonismo, sordera...)

En cas afirmatiu, quin dèficit?

En caso afirmativo, qué déficit?

Patiu algun dèficit motor? Quin?

Padece de algún déficit motor? Cuál?

Heu patit algun accident o malaltia neurològica?
Ha sufrido algún accidente o enfermedad neurológica?

Si us han diagnosticat dislèxia, a quina edat?
Si le han diagnosticado dislexia, a qué edad?

Teniu experiència musical?
Tiene experiencia musical?

- Sí
- No

Des de quina edat i durant quants anys?
Desde qué edad y durante cuántos años?

Quin instrument toqueu?
Qué instrumento toca?

On heu rebut formació musical?
Dónde recibió su formación musical?

- Escola / Escuela
- Institut / Instituto
- Escola de Música / Escuela de Música
- Conservatori / Conservatorio
- Ensenyaments Superiors de Música / Estudios Superiores de Música
- Universitat / Universidad
- Classes particulars / Clases particulares
- Altres:

Quan de temps (en hores i minuts) dediqueu a la música setmanalment?
Cuánto tiempo (en horas y minutos) dedica a la música semanalmente?

Quantes hores al dia escolteu música?
Cuántas horas al día escucha música?

Teniu experiència amb la dansa?
Tiene experiencia con la danza?

En cas afirmatiu, quants anys, mesos o hores li heu dedicat?
En caso afirmativo, cuántos años, meses u horas le ha dedicado?

Indiqueu quin tipus d'esport, competició o activitat física realitzeu o heu realitzat.
Indique qué tipo de deporte, competición o actividad física realiza o ha realizado.

Quantes hores setmanals o mensuals invertiu o invertíeu en exercici físic?
Cuántas horas semanales o mensuales invierte o invertía en ejercicio físico?

Escolteu música durant l'activitat física?
Escucha música durante su actividad física?

- Sempre / Siempre
- Quasi sempre / Casi siempre
- Ocasionalment / Ocasionalmente
- Mai / Nunca

Feu alguna manualitat com dibuixar, pintar, fer ceràmica, cosir, etc.? Quina/es?
Hace alguna manualidad como dibujar, pintar, hacer cerámica, coser, etc.? Cuál(es)?

AVALUACIÓ DE L'EXPERIMENT

EVALUACIÓN DEL EXPERIMENTO

Introdueix les sigles del teu nom i cognoms
Introduce las siglas de tu nombre y apellidos

Com et sents, un cop acabat l'experiment?
Cómo te sientes, una vez terminado el experimento?

Has tingut alguna dificultat per percebre els estímuls? En cas afirmatiu, quina?
Has tenido alguna dificultad para percibir los estímulos? En caso afirmativo, cuál?

Com has percebut l'estímul auditiu?
Cómo has percibido el estímulo auditivo?

▼

Com has percebut l'estímul visual?
Cómo has percibido el estímulo visual?

▼

Quin mètode has utilitzat per agrupar els estímuls mentalment?
Qué método utilizaste para los estímulos mentalmente mentalmente?

Si comptaves, en quin idioma?
Si contabas, en qué idioma?

De l'1 (GENS) al 10 (MOLT), quant consideres que t'has perdut amb els estímuls auditius?
Del 1 (NADA) al 10 (MUCHO), cuánto consideras que te has perdido con los estímulos auditivos?

1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

De l'1 (GENS) al 10 (MOLT), quant consideres que t'has perdut amb els estímuls visuals?
Del 1 (NADA) al 10 (MUCHO), cuánto consideras que te has perdido con los estímulos visuales?

1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

Observacions, suggeriments, comentaris...
Observaciones, sugerencias, comentarios...

4. Generalized Linear Model

Selected data from the questionnaires to apply the GLM

Participant number	Years of music	Dance training
1	22	N0
2	27	YES
3	13	YES
4	16	N0
5	12	N0
6	14	N0
7	18	N0
8	12	YES
9	15	N0
10	17	N0
11	18	N0
12	19	YES
13	12	YES
14	21	YES
15	20	N0
16	13	YES