

Accepted Manuscript

Enhanced brainstem and cortical encoding of sound during synchronized movement

Sylvie Nozaradan, Marc Schönwiesner, Laura Caron-Desrochers, Alexandre Lehmann

PII: S1053-8119(16)30322-6  
DOI: doi: [10.1016/j.neuroimage.2016.07.015](https://doi.org/10.1016/j.neuroimage.2016.07.015)  
Reference: YNIMG 13311

To appear in: *NeuroImage*

Received date: 2 February 2016  
Accepted date: 6 July 2016



Please cite this article as: Nozaradan, Sylvie, Schönwiesner, Marc, Caron-Desrochers, Laura, Lehmann, Alexandre, Enhanced brainstem and cortical encoding of sound during synchronized movement, *NeuroImage* (2016), doi: [10.1016/j.neuroimage.2016.07.015](https://doi.org/10.1016/j.neuroimage.2016.07.015)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**ENHANCED BRAINSTEM AND CORTICAL ENCODING OF SOUND  
DURING SYNCHRONIZED MOVEMENT**

Sylvie Nozaradan<sup>1,2,3,4</sup>, Marc Schönwiesner<sup>3,4,5</sup>, Laura Caron-Desrochers<sup>3,4,5</sup>, Alexandre Lehmann<sup>3,4,5,6</sup>

<sup>1</sup>MARCS Institute, Western Sydney University, Sydney, Australia

<sup>2</sup>Institute of Neuroscience (IONS), Université catholique de Louvain (UCL), Belgium

<sup>3</sup>International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada

<sup>4</sup>Center for Research on Brain, Language and Music (CRBLM), Montreal, Canada

<sup>5</sup>Faculty of Psychology, Université de Montréal, Montreal, Canada

<sup>6</sup>Otolaryngology Department, Faculty of Medicine, McGill University Hospital, Montreal, Canada

Corresponding author:

Dr. Sylvie Nozaradan

Institute of Neurosciences

Université catholique de Louvain

53, Avenue Mounier – UCL 53.75

B-1200 Bruxelles

Belgium

Phone: +32-2-764-9349

Email: sylvie.nozaradan@uclouvain.be

Conflict of interest: The authors declare no competing financial interest.

**ABSTRACT**

Movement to a steady beat has been widely studied as a model of alignment of motor outputs on sensory inputs. However, how the encoding of sensory inputs is shaped during synchronized movements along the sensory pathway remains unknown. To investigate this, we simultaneously recorded brainstem and cortical electro-encephalographic activity while participants listened to periodic amplitude-modulated tones. Participants listened either without moving or while tapping in sync on every second beat. Cortical responses were identified at the envelope modulation rate (beat frequency), whereas brainstem responses were identified at the partials frequencies of the chord and at their modulation by the beat frequency (sidebands). During sensorimotor synchronization, cortical responses at beat frequency were larger than during passive listening. Importantly, brainstem responses were also enhanced, with a selective amplification of the sidebands, in particular at the lower-pitched tone of the chord, and no significant correlation with electromyographic measures at tapping frequency. These findings provide first evidence for an online gain in the cortical and subcortical encoding of sounds during synchronized movement, selective to behavior-relevant sound features. Moreover, the frequency-tagging method to isolate concurrent brainstem and cortical activities even during actual movements appears promising to reveal coordinated processes along the human auditory pathway.

Keywords: EEG, neural entrainment, sensorimotor synchronization, steady-state evoked potentials, frequency-following responses, human auditory brainstem

## 1. INTRODUCTION

Movement synchronization to a steady beat is a fundamental sensorimotor skill, which has been extensively studied as a model of tight temporal alignment of motor outputs to sensory inputs (Repp, 2005; 2012). Most of the studies have focused on the neural correlates of the motor output (see, e.g., Gerloff et al., 1998; Penhune et al., 1998; Jäncke et al., 2000; Ullen et al., 2008; Bengtsson et al., 2009; Hove et al., 2013). However, little is known about the impact of motor outputs on the processing of sensory inputs on which the movement is synchronized. Recent models of active sensing have proposed that the motor system could modulate the cortical processing of sensory information. This modulation would optimize the coordination of motor output on sensory inputs over time through cyclic fluctuations in sensory gain related to motor acts (Morillon et al., 2014; Schubotz, 2007; Schroeder et al., 2010; Zaghera et al., 2013). This dynamic sensory gain could act as a filter, sharpening the temporal representation of auditory inputs, thus facilitating perception of behavior-relevant items (e.g. temporal structure of a dynamic sensory input to which a movement needs to be aligned in a sensorimotor task) (see e.g. Morillon et al., 2014).

Importantly, this online modulation of the sensory input may arise earlier in the auditory hierarchy, namely at brainstem level of the sound encoding. Indeed, a specific link between sensorimotor synchronization behavior and brainstem auditory processing has been highlighted in previous work (Tierney and Kraus, 2013; 2014). For example, it has been shown that the variability in tapping on a steady beat correlates with the variability in response latency in the auditory brainstem (Tierney and Kraus, 2013). Moreover, to account for the fast corrections usually observed in tapping movements synchronized to sound onsets (phase corrections <100 ms of latency in finger tapping on sound sequences; see e.g. Hove et al., 2014), recent work has suggested the existence of a rapid subcortical pathway

involved in movement correction (Schwartz and Kotz, 2013; Hove et al., 2014; Dumas et al., 2005; Bijsterbosch et al., 2011). Specifically, a rapid subcortical/cerebellar pathway (via the brainstem's dorsal cochlear nucleus) may be responsible for the highly accurate encoding of sound onsets timing and the fast transmission of the temporal structure of sounds (i.e. the sound envelope) to thalamic and cortical targets (Schwartz and Kotz, 2013). This rapid feedforward route with highly accurate temporal precision, modulated itself via corticofugal fibers in a continuous feedforward-feedback dynamic, could be critical in sensorimotor synchronization (Schwartz and Kotz, 2013), thus placing the brainstem auditory nuclei in a key role in monitoring sensorimotor temporal alignment. However, there is currently no direct evidence of the modulation of the auditory brainstem response when synchronizing movements to sounds.

The general goal of our study was to fill this gap in knowledge by directly testing whether cortical but also brainstem responses to sound are shaped during synchronized movement to the sound. Previous work has provided evidence that brainstem auditory processing can be dynamically modulated by task contexts in humans (Galbraith et al., 1998; 2003; Rinne et al., 2008; Chandrasekaran et al., 2009; Lehmann and Schönwiesner, 2014) and also non-human animals (Slee and David, 2015). This sensitivity to ongoing contextual and task demands has been explained by the existence of efferent connections from the cortex to subcortical nuclei (see e.g. Winer, 2006; Winer and Lee, 2007; Bajo and King, 2013). These efferent corticofugal fibers project to the major levels of the auditory system, including the thalamus, the inferior colliculus, the superior olivary complex and the cochlear nucleus, allowing the brainstem encoding of sounds to be modulated depending on the level of activity in cortical auditory areas (Diamond et al., 1969; Weedman and Ryugo, 1996; Winer et al., 1998; Mulders and Robertson, 2000).

Building on this work, the current study aimed to test a more specific hypothesis, - i.e. that the brainstem encoding of sound would be enhanced during auditory-motor synchronization. More specifically, we hypothesized that this sensory gain would be selective to the sound envelope, as this sound component drives the movement synchronization, in contrast with the tones' pitch which is a sound component not directly relevant to the tapping behavior. Hence, we aimed to test whether the envelope representation at brainstem level was enhanced *during* synchronized movements, thus providing first evidence of auditory brainstem modulation driven by a task involving sensorimotor temporal alignment. Alternative to this sensory gain hypothesis however, one could predict to observe a sensory attenuation, due to an effect of agency related to the synchronization between sound and tapping (Schröger et al., 2015; Timm et al., 2014).

Processes of sensory modulation throughout the human auditory pathway remain poorly understood, probably due to the scarcity of methods to simultaneously measure the activity at different levels of the auditory pathway with high temporal resolution, specifically while performing movements. The current study addressed this question by investigating concurrent activities produced by distant neural populations throughout the auditory hierarchy. Participants listened to a chord made of three tones periodically amplitude-modulated in order to induce a steady beat. They were asked to carefully listen either without moving or to tap in synchrony to every second beat. Importantly, our study made an original use of the frequency-tagging method to simultaneously measure cortical and brainstem responses to sounds with electroencephalography (EEG). The frequency-tagging method consists in identifying and isolating different components of the neural response to a stimulus based on the expected frequencies of these components (i.e. steady-state evoked potentials, or SS-EPs; see e.g. Galambos et al., 1981; Rees et al., 1986; Picton et al., 1987;

Regan, 1989; Ross et al., 2000; 2004; Bidet-Caulet et al., 2007; Norcia et al., 2015). In the current study, these frequencies were determined by the frequency structure of the stimulus itself (i.e. sound envelope at 2.4 Hz and harmonics, partials at 200, 400 and 600 Hz and sidebands at partial frequencies  $\pm$  2.4 Hz and harmonics). Importantly, the frequency content of the sound was specifically set such as to elicit frequency-locked responses that are likely to originate in most part from the brainstem auditory nuclei (i.e. partials and sidebands > 200 Hz) and the cortex (i.e. amplitude modulation < 3 Hz) respectively (Skoe and Kraus, 2010). Hence, the frequency-tagging approach appears to be particularly well suited to isolate concurrent cortical and brainstem processes.

We also moved the method a step further by capturing these activities not merely during perception but also during actual movements synchronized on the dynamic sensory input, thus providing first empirical cues about the possible effects related to auditory-motor coordination on the subcortical encoding of sound. Moreover, a significant signal-to-noise ratio was expected for both brainstem and cortical activities, although the auditory stimulation lasted less than 7 min in total per condition. Indeed, the use of long-lasting continuous stimuli allowed a fine spectral resolution (0.02 Hz) to be obtained in the current study, thus concentrating the frequency-locked activities within very narrow frequency bands in the EEG spectrum (Norcia et al., 2015). In addition, the use of long-lasting sequences of continuous stimuli were likely to improve the opportunity to capture dynamic aspects of the processes at stake in sensorimotor synchronization, in contrast with transient event-related potentials (ERPs) reflecting transient neural responses triggered by the occurrence of transient stimuli (Norcia et al., 2015; Nozaradan, 2014).

## 2. MATERIALS AND METHODS

### 2.1. Participants

Twenty-two healthy volunteers (10 females, 12 males, all right-handed, mean age  $21.2 \pm 2.9$  years, aged between 18 and 29) took part in the study after providing written informed consent. They were either music amateurs as listeners or dancers, or considered as musicians according to the number of years of music practice (11 participants with  $8.5 \pm 3.8$  years). None had prior experience with the tapping task used in the present study. They had no history of hearing, neurological or psychiatric disorder, and were not taking any drug at the time of the experiment. The study was approved by the Research Ethics Committee of the Faculty for Arts and Sciences of the University of Montreal.

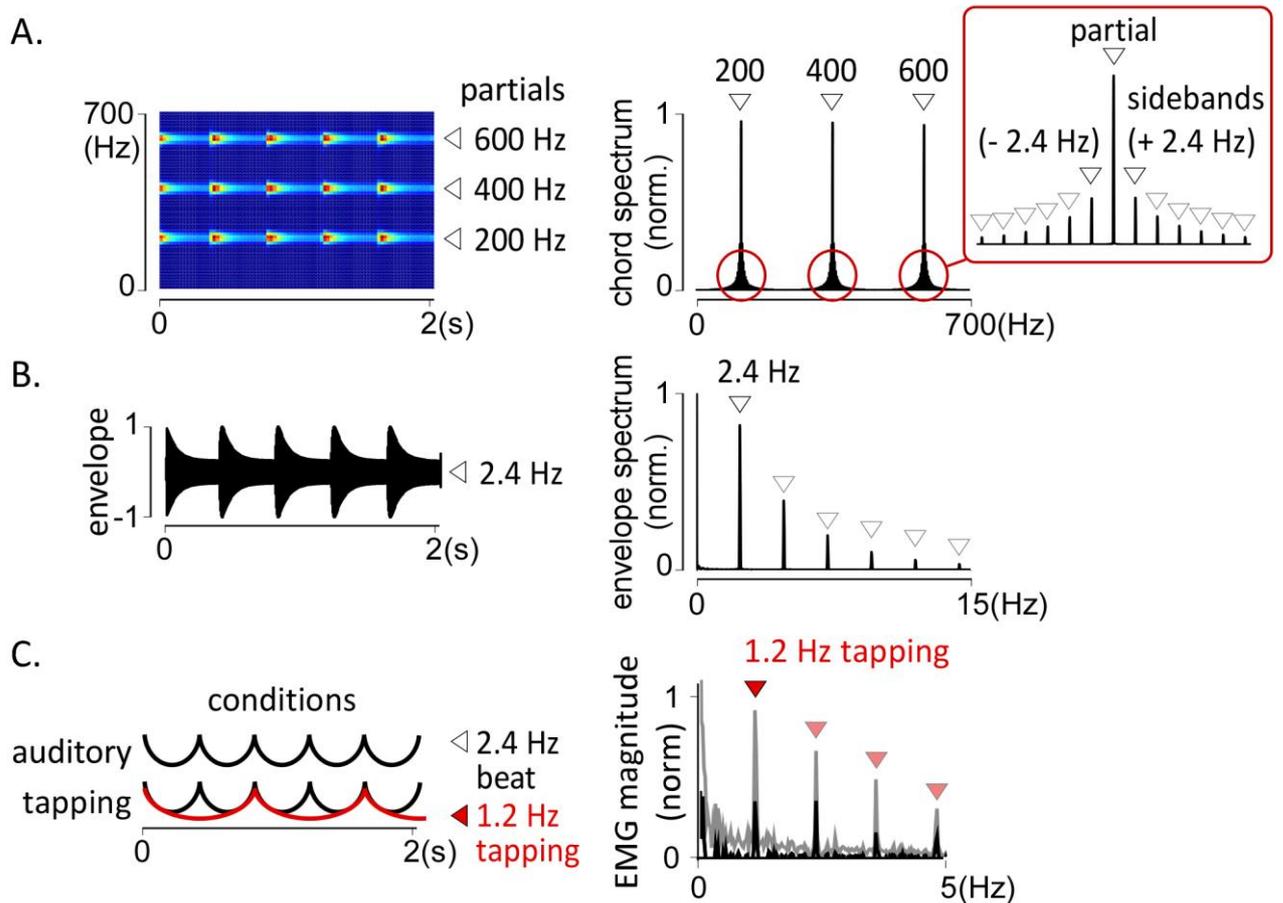
### 2.2. Auditory stimulation

The auditory stimulus consisted in a continuous sequence of 39.5 s. The sound was composed of a harmonic tone with three partials at 200, 400, and 600 Hz composing a consonant chord (Fig. 1). These frequencies were chosen because frequency-following (i.e. frequency-locked) responses at these rates recorded in humans with scalp EEG are most likely to originate from brainstem nuclei of the ascending auditory pathway due to the low-pass limitation in the production of sustained frequency-locked responses to sounds in the human auditory cortex (Skoe and Kraus, 2010). A harmonic relationship between the three tones was chosen to optimize the emergence of brainstem frequency-following responses to the chord, as brainstem responses to consonant intervals are known to be more robust than for non harmonic intervals (Bidelman, 2013). A beat was induced by modulating the amplitude of the chord at a frequency of 2.4 Hz (144 beats per minute), using an asymmetrical Hanning envelope (12 ms rise time and 404 ms fall time with a modulation depth of 75%, as in Nozaradan et al., 2015). A 2.4 Hz frequency was chosen because (i) this

tempo lies within the ecological range of beat perception and production (Drake and Botte, 1993), (ii) we previously showed that this beat frequency elicits a measurable steady-state evoked potential (SS-EP) in the human EEG (Nozaradan et al., 2011; 2015), (iii) pilot experiments showed that participants are comfortable tapping on every second beat at this beat frequency, and (iv) frequency-locked responses to amplitude modulation of a carrier tone below 20 Hz are most likely to originate from cortical auditory sources rather than brainstem nuclei when recorded with scalp EEG (Wong and Stapells, 2004; Ross et al., 2000). An asymmetrical Hanning window was chosen because (i) it is well suited to tapping synchronization due to the sharp slope of the rising ramps, and (ii) it was previously shown to elicit a measurable cortical beat SS-EP (Nozaradan et al., 2011; 2015). Stimuli were generated using Matlab (The MathWorks, Natick, USA) and presented binaurally through electromagnetically shielded insert earphones at 75 dB SPL (ER 2, Etymotic Research, Elk Grove Village, IL, USA).

To determine the frequencies at which the brainstem and cortical SS-EPs were expected in the EEG, the auditory stimulus was transformed in the frequency domain using a discrete Fourier transform. The frequencies of interest appeared in the obtained spectrum as three peaks of equal acoustic energy at 200, 400 and 600 Hz, flanked by sidebands at  $[+/-] 2.4$  Hz [and multiples] above and below each partial, corresponding to each partial periodically amplitude-modulated at beat frequency by the non-sinusoidal envelope (Fig. 1).

To determine the frequencies at which the cortical frequency-locked response to the envelope was expected in the EEG, the temporal envelope of the auditory stimulus was transformed into the frequency domain using a discrete Fourier transform. The frequencies of interest appeared in the obtained spectrum as peaks of acoustic energy at 2.4 Hz and harmonics (Fig.1).



**Figure 1. Experimental design and stimulus.** A. Time-frequency excerpt of the stimulus. The harmonic chord (200, 400 and 600 Hz) was amplitude-modulated at 2.4 Hz to induce a steady beat. The spectrum of this sound exhibits three peaks at the partials' frequencies, each flanked by sidebands at  $\pm 2.4$  Hz and harmonics. B. Excerpt of the sound envelope with the periodic non-sinusoidal amplitude modulation. The envelope spectrum presents peaks at 2.4 Hz and harmonics. C. In the auditory condition, participants carefully listened to the sound without moving. In the tapping condition, they were asked to listen to the same sound and tap the right hand in synchrony on every second beat (at 1.2 Hz, i.e. at half the beat frequency). On the right, the spectrum represents the group-level average (black) and standard deviation (grey) of the electromyographic (EMG) signal (magnitude normalized between 0 and 1) recorded on the right forearm during the tapping. As expected, a peak is observed at 1.2 Hz and harmonics.

### 2.3. Experimental conditions

Participants were comfortably seated in a chair with their head resting on a support. They were instructed to relax, avoid any unnecessary head or body movement and keep their eyes fixed on a point displayed on the wall in front of them. They were asked to perform two

different tasks: first, an auditory task, and second, a tapping task, in separate runs (Fig. 1C). Each run consisted of 10 trials, each starting with 3 s of silence, followed by the 39.5-s sound stimulus. The polarity of the sound waveform was inverted on every other trial to prevent potential electrical artifacts at the frequencies of the sound input (Skoie and Kraus, 2010). However, our insert earphones were electrically shielded and no such artifacts were observed at either polarity. The experiment always started with the auditory task. Each trial onset was self-paced. To ensure that participants focused their attention on the sound in the auditory task, they were asked to carefully listen to the sound and report any irregularity in the sound and specifically in the beat duration at the end of each trial. There was no actual irregularities in any of the sequences (so that there was no extensive assessment of the detection performance), but probably due to the long-lasting repetition of the sound participants reported hearing subtle changes throughout the trial in half of the trials on average.

During the second condition, participants were asked to perform a tapping with their right hand as accurately as possible on every other beat of the sequence, i.e., at half the beat frequency ( $f_{\text{beat}}/2 = 1.2$  Hz) (Fig. 1). This design aimed to reduce the frequency overlap between the responses to the sound envelope (i.e. the beat at 2.4 Hz and harmonics) and the neural activity related to motor activity (i.e. the tapping at 1.2 Hz and harmonics), although an overlap of the two activities at 2.4 Hz cannot be excluded completely (Nozaradan et al., 2015). All participants underwent a short training session before the experiment to ensure that they understood the task. They were asked to start tapping on the third auditory beat of each stimulus, and to maintain their movement as synchronized as possible throughout the entire trial. The tapping was performed with small up and down movements of the hand starting from the wrist joint, maintaining the forearm and elbow

fixed on an armrest cushion. When performing the tapping movement, the fingers of the tapping hand came transiently in contact with the armrest cushion. All participants naturally synchronized their movement such that the occurrence of this contact coincided with the occurrence of the beat. Importantly, the contact with the armrest produced tactile feedback, but did not produce any auditory feedback, as the participants were fitted with earphone inserts. The experimenter regularly went to the recording room, to monitor compliance to these instructions.

#### **2.4. EEG recording**

The EEG was recorded using 64 active sintered Ag-AgCl electrodes placed on the scalp according to the International 10/10 system (ActiveTwo, BioSemi, The Netherlands). Active electrodes contain the first amplifier stage within the electrode cover and provide impedance transformation on the electrode to prevent interference currents from generating significant impedance-dependent nuisance voltages. We therefore did not control electrode impedances but rather kept direct-current offset close to zero during electrode placement. Vertical and horizontal eye movements were monitored using three additional electrodes placed on the outer canthus of each eye and on the inferior area of the left orbit. Two additional electrodes were placed on the right forearm (on the extensor carpi ulnaris) to record the electromyographic activity generated by the hand tapping in the movement condition. Although this montage does not provide an extensive recording of the electromyographic activity corresponding to the flexor and extensor muscles or an accurate measure of the tapping movement starting points, it provides an estimation of the frequency of the movement throughout the trials, thus allowing the compliance of the participants to the instruction to tap at 1.2 Hz to be indexed. Reference-free electrode signals were amplified, sampled at 8192 Hz (ActiveTwo amplifier, BioSemi, The Netherlands),

and stored for offline analysis using BioSemi ActiView software. To facilitate computer processing of the recorded data, the signal of three EEG channels (Fz, FCz and Cz) was exported at 8192 Hz for further analysis, as these electrodes are known to capture the major part of auditory brainstem and cortical responses (Skoe and Kraus, 2010). Those signals were re-referenced to average mastoids, a standard montage to measure brainstem and cortical auditory responses (Skoe and Kraus, 2010).

### **2.5. SS-EPs analysis**

EEG processing was carried out using Letswave5 ([www.nocions.org/letswave5](http://www.nocions.org/letswave5)) running under Matlab (The MathWorks, Natick, USA). Continuous EEG recordings were filtered using a 0.1-Hz high-pass fast-Fourier transform (FFT) filter to remove slow drifts in the recorded signals. Epochs lasting 39.5 s were obtained by segmenting the recordings from +0 to +39.5 s relative to the onset of the auditory stimulus, thus yielding 10 epochs for each participant and condition. Baseline correction of these epochs was computed by subtracting the amplitude of each epoch by the average amplitude measured from -0.6 to -0.2 s relative to the onset of the auditory stimulus. A notch FFT filter centered on 60 Hz with a width of 1 Hz was used to remove the artefact due to the electric line noise.

For each participant and condition, EEG epochs were averaged across trials. The time-domain averaging procedure was used to enhance the signal-to-noise ratio of the EEG activities elicited in response to the stimulus by attenuating the contribution of activities that were not strictly phase-locked to the sound stimulus across trials. The obtained average waveforms were then transformed into the frequency domain using a discrete Fourier transform, thus yielding spectra with 0.02 Hz resolution. These spectra were normalized by subtracting the average amplitude measured at -3 to -5 and +3 to +5 frequency bins relative to each frequency bin across the x-axis of the spectra. Hence, in the absence of an SS-EP, the

signal amplitude at a given frequency should be similar to the signal amplitude of the mean of the surrounding frequency bins (Mouraux et al. 2011, Nozaradan et al. 2011; 2012a; 2012b; 2015; Chemin et al., 2014; Nozaradan et al., 2016). Importantly, while alternative approaches might be used to calculate the signal-to-noise ratio of the responses, the subtraction procedure used here did not specifically aim to calculate a signal-to-noise ratio of the responses. Rather, it was used to correct for the background noise when measuring the amplitude of the responses over the different frequency ranges, thus being closer to procedures of spectrum baseline flattening through running subtraction (see e.g. Kay, 1988; Wang et al., 2012). Moreover, it allows the amplitudes to be provided in microvolt units, thus giving an intuitive idea on the actual size of the responses throughout the frequency spectrum. For each condition and participant, the obtained spectra were then averaged across the three scalp electrodes. The magnitude of the SS-EPs was then estimated by taking the maximum noise-subtracted amplitude measured in a range of 3 frequency bins centered over the expected SS-EP frequency, based on the spectrum of the sound envelope. This range of frequencies allowed accounting for possible spectral leakage due to the occurrence of small fluctuations of the form of the response over the cycles throughout the sequences. The magnitude of the cortical responses to the beat, further referred to as *cortical SS-EP*, was estimated at the frequency of the beat and its harmonics (to the sixth one, thus 2.4, 4.8, 7.2, 9.6, 12 and 14.4 Hz). Indeed, given the non-sinusoidal waveform of the beat amplitude-modulation applied to the chord, a number of harmonics were expected to be elicited (Fig. 1). Moreover, the magnitude of the brainstem SS-EPs was estimated at the frequency of the three partials of the chord (200, 400 and 600 Hz), further referred to as *brainstem partials SS-EPs*. Finally, the magnitude of the brainstem response to the amplitude-modulation of the sound, further referred to as *brainstem sidebands SS-EPs*, was obtained at the frequencies of

the sidebands flanking each partials, thus at + and - 2.4 Hz to the sixth harmonics (as for the cortical beat SS-EP) (Fig. 1).

## 2.6. Statistical evaluation

Statistical analyses were performed using SPSS Statistics 21.0 (IBM, Armonk, NY, USA). Significance level was set at  $p < 0.05$ . When relevant, the Greenhouse-Geisser correction was used to correct for violations of sphericity in the performed ANOVAs.

We first examined whether the auditory stimulus elicited cortical SS-EPs and brainstem partials and sidebands SS-EPs at the expected frequencies in the two conditions. For each condition and participant, a one-sample t-test was used to determine whether the amplitudes of the SS-EPs were significantly different from zero (Table 1). Indeed, in the absence of significant entrainment to the input, the averaged noise-subtracted signal amplitude may be expected to tend towards zero (Mouraux et al., 2011).

The general goal of the study, however, was to test the hypothesis of a boost of amplitude of the SS-EPs in the tapping compared to the auditory condition. This hypothesis was tested using paired t-tests comparing the amplitudes across the two conditions for the cortical SS-EPs and the brainstem partials and sidebands SS-EPs.

Our second goal was then to specify how the gain in amplitude in the tapping condition compared to the auditory condition was distributed over the different components of the brainstem response to the sound, i.e. over the sideband and partial SS-EPs and over the three tones of the chord. The gain in amplitude was calculated for each participant as the percentage of difference in amplitude between the tapping minus the auditory condition. Hence, in the absence of significant gain in amplitude in the tapping condition, the percentage may be expected to tend towards zero. Most importantly, this normalization allowed the distribution of the gain to be compared across the different components of the

brainstem SS-EPs, irrespective of the expected difference in absolute amplitude between these components. Indeed, independently of the conditions, a brainstem frequency-following response is expected to present a greater amplitude at 200 Hz compared to 400 and 600 Hz, as the frequency-locking of auditory brainstem neural populations becomes weaker with increasing frequency (Greenberg et al., 1987; Krishnan, 2007), reflecting the low-pass nature of brainstem frequency-locking (Skoie and Kraus, 2010). Moreover, according to the spectral structure of the sound itself, the brainstem responses to the sidebands are expected to be much smaller than their respective partials (Fig. 1). However, independently of this amplitude difference, the distribution of the *gain* in amplitude when tapping could be hypothesized not to be uniform over the different components of the brainstem response. The obtained measures of gain were thus compared using a 2 x 3 repeated-measures ANOVA with the factors 'component' (two levels, for the partial and sideband SS-EPs) and 'frequency' (three levels, for the three tones of the chord). A significant interaction between the two factors would indicate that the different components of the brainstem response to the sound did not present similar profiles of gain during the tapping condition. When significant, post hoc comparisons were conducted using paired t-tests.

Moreover, a correlation between the gain of the cortical SS-EPs and the different components of the brainstem SS-EPs was tested across participants. A multiple linear regression analysis was conducted with the cortical gain as dependent variable and the gain obtained for the three partials and the three sidebands SS-EPs as independent variables. For all brainstem SS-EP components, partial Pearson's correlations were then used to examine the association between cortical and brainstem gain in amplitude when tapping.

Finally, we also tested the hypothesis of a direct relationship between the tapping movement precision, as estimated by the amplitude of the peak obtained at the movement frequency (1.2 Hz and harmonics) in the EMG signal, and the cortical and brainstem auditory gain. A multiple linear regression analysis was conducted across participants with the EMG value as dependent variable and the sensory gain obtained for the cortical and brainstem partials and sidebands SS-EPs as independent variables. For all SS-EP components, partial Pearson's correlations were then used to examine the association between the EMG value and the sensory gain when tapping.

*Phase lag of cortical and brainstem SS-EPs.* Cortical and brainstem responses to the sound envelope were hypothesized to be generated by cortical and brainstem nuclei of the ascending auditory pathway respectively. If so, these two responses should be related not merely in frequency (both frequency-locked at 2.4 Hz beat frequency) but also in phase (i.e. elicited with a consistent lag due to the distinct latencies of these responses if they actually originate from distinct nuclei of the auditory hierarchy). Phase values of the cortical response to the sound envelope were obtained for each participant and condition at 2.4 Hz from the complex-valued Fourier transform of the epochs averaged across the three channels. To obtain the phase values of the brainstem responses to the sound envelope, the brainstem responses to the sound envelope were first isolated for each participant and condition by filtering each epoch using a 190 Hz - 610 Hz bandpass FFT filter (thus isolating the portion of signal corresponding to the brainstem response to the sound) and then extracting the envelope of this response using a Hilbert transform as implemented in Matlab. The phase values of these brainstem responses to the envelope were finally measured from the complex-valued Fourier transform of these envelopes averaged across the three channels, as for the cortical SS-EP phase estimation.

A  $2 \times 2$  repeated-measures circular ANOVA as implemented in a Matlab toolbox (Berens, 2009) was used to compare these values, with the response component (cortical and brainstem SS-EP) and the conditions (auditory and tapping condition) as factors. Importantly, the aim of this analysis was not to provide an absolute measure of latency of the expected cortical and brainstem activity frequency-locked at 2.4 Hz, but to provide an estimate of the phase lag between the two responses (see e.g. Ross et al., 2000).

*Time course of cortical and brainstem SS-EPs.* The time course of the responses to the sound envelope was also analyzed for each participant and condition, to investigate the shape of the responses and to test whether it was modulated by the tapping. The time course of cortical responses to the sound envelope was obtained by filtering the EEG signal between 0.3 and 30 Hz (FFT filter). The time course of brainstem responses to the sound envelope was obtained by filtering the EEG signal between 190 and 610 Hz (FFT filter), then by extracting the envelope of the response using a Hilbert transform.

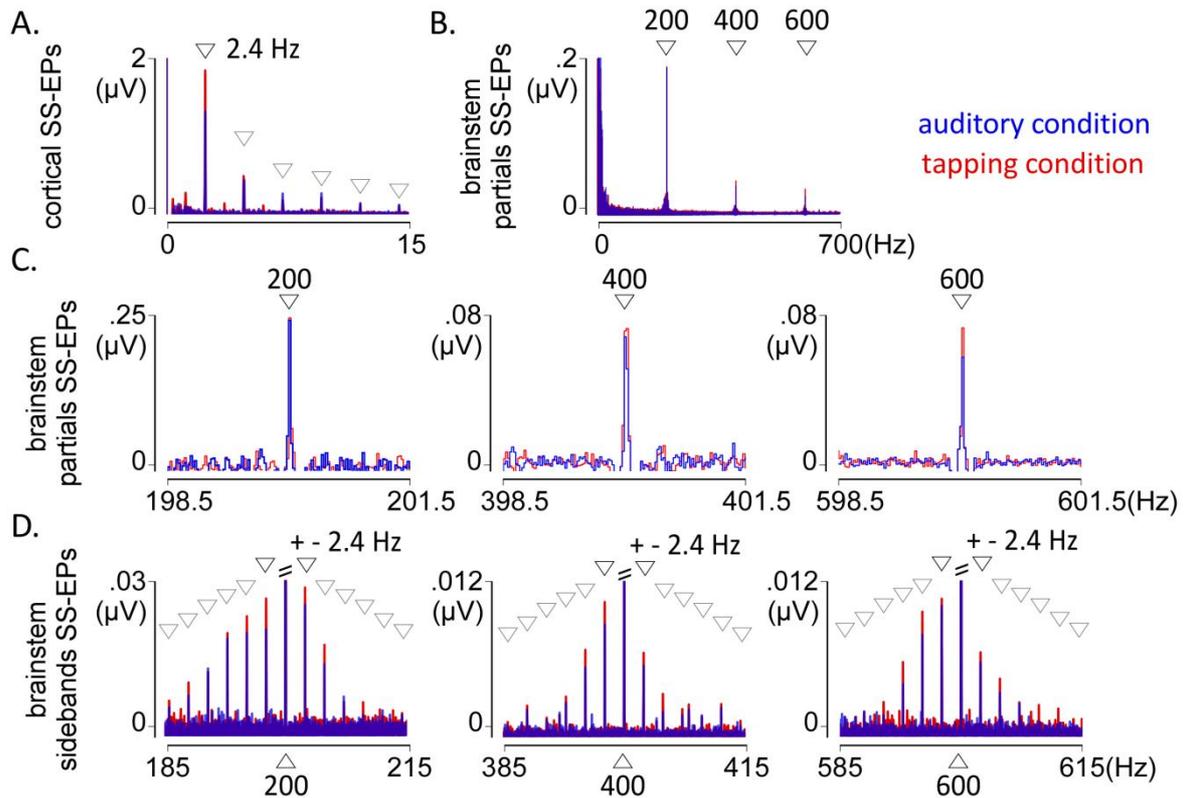
For both cortical and brainstem data, epochs were then segmented into chunks of 416 ms long, corresponding to the window between each beat onset (starting at + 832 ms relative to the onset of the entire sound sequence, to avoid the evoked potentials elicited by the onset of the sound sequence). Furthermore, the 416-ms chunks were split into odd and even beats, allowing the comparison of the responses to the beats with and without synchronized tapping (odd and even beats respectively).

Because there was no a priori assumption regarding the latencies at which a difference could arise between these waveforms, we used a point-by-point comparison by means of a cluster-based permutation approach (Maris and Oostenveld, 2007; Van den Broeke and Mouraux, 2015). The technique assumes that true neural activity will tend to generate signal changes over contiguous time points (Groppe et al., 2011). First, the waveforms were

compared by means of a point-by-point paired t-test. Then, clusters of contiguous time points above the critical t value for a parametric two-sided test were identified, and an estimate of the magnitude of each cluster was obtained by computing the sum of the t-values constituting each cluster (cluster-level statistics). Random permutation testing (1000 times) of the subject-specific waveforms of the different chunks (performed independently for every subject) was then used to obtain a reference distribution of maximum cluster magnitude. Finally, the proportion of random partitions that resulted in a larger cluster-level statistic than the observed one (i.e. p value) was calculated. Clusters in the observed data were regarded as significant if they had a magnitude exceeding the threshold of the 97.5th and 2.5th percentiles.

### 3. RESULTS

As shown at the group-level average of the frequency spectra (Fig. 2), the auditory input elicited significant cortical SS-EPs at 2.4 Hz and harmonics in the two conditions, thus corresponding to responses frequency-locked to the sound envelope. Moreover, significant SS-EPs were observed at 200, 400 and 600 Hz  $\pm$  2.4 Hz and harmonics in the two conditions, thus corresponding to brainstem responses frequency-locked to the three tones and their beat periodic amplitude modulation (see Table 1). This result corroborates our hypothesis to obtain valid estimates of concurrent cortical and brainstem frequency-locked responses to the sound, even when performing movement, with less than 7 min of stimulation per condition.

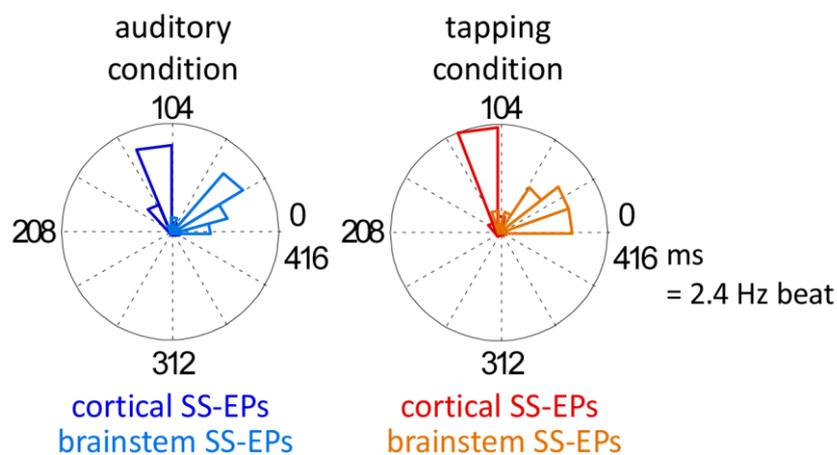


**Figure 2. Spectra of the cortical and brainstem responses.** Average across channels Fz, FCz and Cz and participants, for the auditory (blue) and tapping (red) conditions. A. Cortical SS-EPs at beat frequency (2.4 Hz) and harmonics in response to the amplitude modulation of the sound. B. Brainstem partial SS-EPs at 200, 400 and 600 Hz. C. Spectra zoomed on each of these brainstem partial SS-EPs. D. Brainstem sideband SS-EPs at  $\pm 2.4$  Hz and harmonics in response to the amplitude modulation of the sound. Note that for the three types of response (cortical, brainstem partials, brainstem sidebands), the amplitude in the tapping condition was enhanced as compared to the auditory condition.

		auditory condition	tapping condition	gain
cortical beat SS-EP	<b>2.4 Hz (and harmonics)</b>	t=11.28 ****	t=11.34****	60±63% t=4.50***
brainstem partials SS-EP	<b>200 Hz</b>	t=11.24 ***	t=11.52 ****	12±28% t=2.04 <i>ns</i>
	<b>400 Hz</b>	t=6.31 ****	t=6.51 ****	11±26% t=2.07 <i>ns</i>
	<b>600 Hz</b>	t=4.11 **	t=3.86 **	32±64% t=2.38*
brainstem sidebands SS-EP	<b>200 Hz ± 2.4 H (and harmonics)</b>	t=11.24 ***	t=11.52 ****	54±74% t=3.44**
	<b>400 Hz ± 2.4 Hz (and harmonics)</b>	t=11.24 ***	t=11.52 ****	12±53% t=1.09 <i>ns</i>
	<b>600 Hz ± 2.4 Hz (and harmonics)</b>	t=11.24 ***	t=11.52 ****	29±35% t=3.91***

**Table 1. T-tests against zero of the peaks amplitudes and percentage of gain across conditions.** In the two conditions, the chord elicited significant responses identified in the EEG spectra in the form of brainstem partial steady-state evoked potentials (SS-EPs). Significant responses to the envelope were also elicited in the form of cortical SS-EPs at 2.4 Hz and harmonics and in the form of brainstem sideband SS-EPs. Finally, the gain in amplitude during the tapping (measured as the percentage of difference in amplitude between the tapping and the auditory condition) was significant for the cortical SS-EP but this was not systematically the case for the different brainstem components (mean ± standard deviation; t-test against zero; all df=21). *ns*: non significant; \*:  $p \leq .05$ ; \*\*:  $p \leq .01$ ; \*\*\*:  $p \leq .001$ ; \*\*\*\*:  $p \leq .0001$ .

The phase values obtained for the cortical and brainstem envelope SS-EPs in the auditory and tapping conditions respectively were compared using a 2 x 2 circular ANOVA (response component and condition as factors). This test revealed a significant main effect of response component (cortical vs. brainstem;  $F=147.99$ ,  $p < 0.0001$ ,  $\eta^2=19.01$ ). However, it did not reveal any effect of condition ( $F=0.34$ ,  $p=0.55$ ,  $\eta^2=0.04$ ) nor interaction between the two factors ( $F=0.31$ ,  $p=0.55$ ,  $\eta^2=0.04$ ). This result indicates that the phase lag between the brainstem and cortical responses was significant and consistent across participants, and was not modulated by the experimental conditions (Fig. 3).

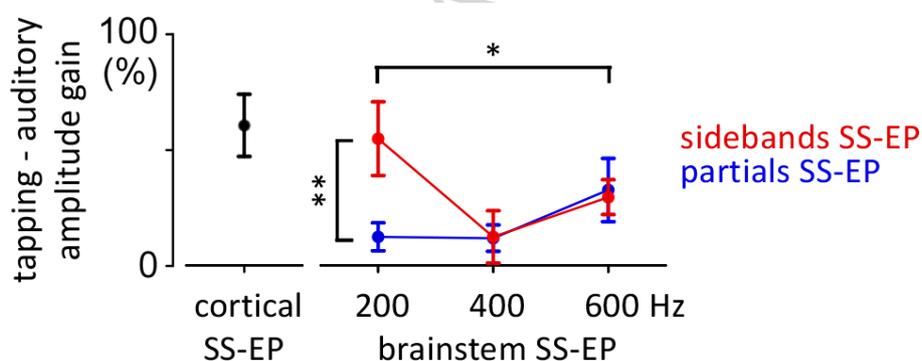


**Figure 3. Phase lag between the brainstem and cortical responses to the sound envelope (at 2.4 Hz).** In the two conditions (blue: auditory, red: tapping condition), there was a consistent phase lag between the two responses, compatible with the hypothesis of a distinct source (cortical vs. brainstem) for these responses to the envelope, as recorded from the scalp.

Of primary relevance to our hypotheses, however, was the amplitude of SS-EPs in the tapping as compared to the auditory condition. The amplitude of the cortical SS-EPs was significantly greater in the tapping condition, as revealed by a paired t-test across conditions ( $t_{21}=4.34$ ,  $p=0.0003$ ). Moreover, the amplitude of the brainstem SS-EPs averaged across the different components was also significantly enhanced in the tapping condition ( $t_{21}=2.45$ ,  $p=0.02$ ), for both the averaged partial SS-EPs ( $t_{21}=2.27$ ,  $p=0.03$ ) and the averaged sideband SS-EPs ( $t_{21}=2.29$ ,  $p=0.03$ ). These results thus corroborate our first hypothesis of a boost in amplitude of the sound encoding when tapping as compared to passive listening.

In order to specify the distribution of this gain in amplitude over the different components of the brainstem response to the sound, a  $2 \times 3$  ANOVA revealed no significant main effect of 'component' (partials vs. sidebands;  $F_{(1,21)}=1.95$ ,  $p=0.17$ ,  $\eta^2=0.08$ ), no significant main effect of 'frequency' (200 vs. 400 vs. 600 Hz;  $F_{(1,78,37.39)}=2.35$ ,  $p=0.11$ ,  $\eta^2=0.10$ ), but a significant interaction between the two factors ( $F_{(1,95,40.97)}=4.80$ ,  $p=0.01$ ,  $\eta^2=0.18$ ) (Fig. 4). Post hoc t-tests revealed a significant difference in the gain between the partials and sidebands at 200

Hz, in favor of the sidebands ( $t_{21}=3.36$ ,  $p=0.003$ ), whereas no significant differences were observed between the partials and sidebands at 400 and 600 Hz ( $t_{21}=0.04$ ,  $p=0.96$  and  $t_{21}=0.22$ ,  $p=0.82$  respectively). These results are also corroborated by the t-tests against zero assessing whether the gain values for the different components of the brainstem response were significant (see Table 1 for detailed values). Together, these results indicate a dissociated profile of amplitude gain in favor of the sidebands, particularly at the lower frequencies of the chord.



**Figure 4. Gain in amplitude in the tapping vs. auditory condition.** Both brainstem and cortical SS-EPs were generally enhanced in the tapping condition. However, this gain in the brainstem response was not uniformly distributed over the partials and sidebands and over the three tones of the chord (interaction between the two factors). Rather, it was relatively more pronounced in the sidebands, in particular at the lower tone of the chord (200 Hz) which was also significantly correlated with the cortical gain ( $p=0.01$ ). \*:  $p<0.05$ , \*\*:  $p<0.01$ .

Moreover, the multiple linear regression analysis with the cortical gain as dependent variable and the gain obtained for the three brainstem partials SS-EPs at 200, 400 and 600 Hz as independent variables did not yield a significant model ( $R^2=0.14$ ,  $F_{3,18}=1.04$ ,  $p=0.39$ ), neither the partial correlations between these values were significant (all with  $p>0.05$ ). In contrast, when taking the gain obtained for the brainstem *sidebands* SS-EPs at 200, 400 and 600 Hz  $\pm$  2.4 Hz and harmonics as independent variables, the model was marginally

significant ( $R^2=0.34$ ,  $F_{3,18}=3.07$ ,  $p=0.054$ ), with significant partial correlations between the cortical and brainstem sidebands' gain at 200 Hz ( $R=0.48$ ,  $p=0.01$ ) and 600 Hz ( $R=0.41$ ,  $p=0.02$ ) but no significant correlation at 400 Hz ( $R=0.15$ ,  $p=0.24$ ). Hence, although the results of these correlation tests did not yield robust statistical power, they suggest a particular relationship between the cortical gain and the sidebands, in particular at the lower tone of the chord, in contrast with the partial responses.

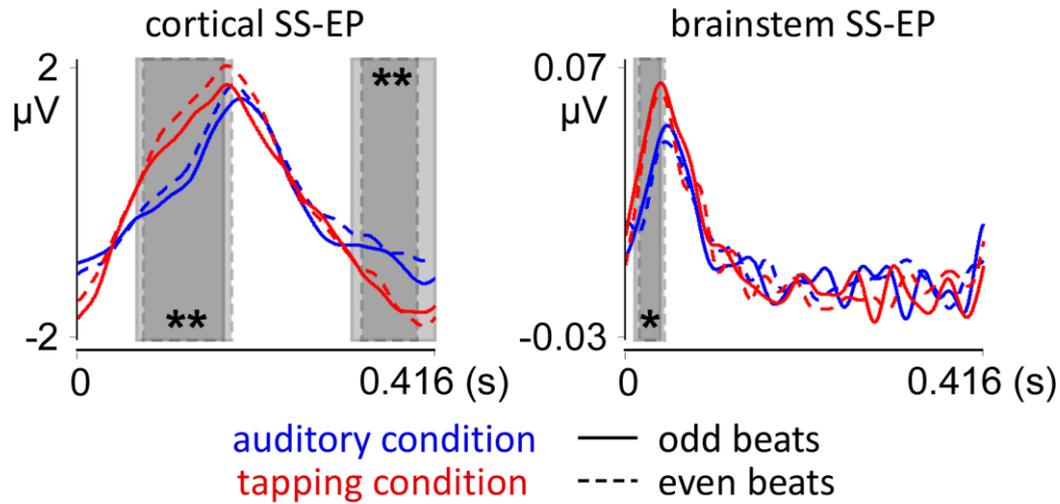
We also examined whether a movement on every second beat yielded an enhanced amplitude of the response of every second beat in the tapping condition. If so, this should correspond to the emergence of a peak at 1.2 Hz and harmonics, and at 200, 400 and 600 Hz  $\pm$  1.2 Hz and harmonics in the EEG spectra. The amplitude of the cortical activity at 1.2 Hz and harmonics f3 and f5 (i.e. at 3.6 Hz and 6 Hz, thus not including the harmonics which overlap with 2.4 Hz and harmonics) were measured from the EEG spectra of each participant in the two conditions. Amplitudes of  $0.16 \pm 0.18$   $\mu$ V (mean and standard deviation) were obtained in the tapping condition at 1.2 Hz and harmonics, which was significantly greater than zero (one-sample t-test:  $t_{21}=4.03$ ,  $p=0.0006$ ). In contrast, there was no significant activity at these frequencies in the auditory condition ( $-0.01 \pm 0.08$ ,  $t_{21}=0.68$ ,  $p=0.50$ ). The amplitude of the brainstem activity at 200, 400 and 600 Hz  $\pm$  1.2 Hz and harmonics were also analyzed for each participant in the two conditions. None of these measures were significantly different from zero, in any of the two conditions (all  $p>0.1$ ). These results indicate that while the cortical SS-EPs were modulated at a frequency corresponding to the movement, this was not the case for the brainstem responses.

Finally, the multiple linear regression analysis with the EMG value as dependent variable and the gain obtained for the cortical and brainstem SS-EPs as independent variables did not yield a significant model ( $R^2=0.02$ ,  $F_{3,18}=0.13$ ,  $p=0.94$ ). None of the partial correlations

between these values were significant (all with  $p > 0.21$ ). This absence of significant correlation may be because there was no actual correlation, thus suggesting an indirect relationship between the observed sensory gain and the actual tapping movement, or may be because there was not enough variance in the data (participants were not specifically selected in the current study to maximize interindividual differences in tapping accuracy).

*Time course of cortical and brainstem SS-EPs.* The time course of cortical and brainstem responses was compared across the auditory and tapping conditions, yielding observations which corroborated the results obtained using the frequency domain analysis. Specifically, the cortical responses were significantly increased in the tapping condition as compared to the auditory condition, in a temporal window spanning 71 to 169 ms (averaged p value over the window = 0.01) and between 320 and 416 ms (average p value = 0.01) for the even beats (Fig. 5). Similarly, the cortical responses to the odd beats also significantly increased in the tapping condition, between 73 and 175 ms (average p value = 0.01) and between 332 and 396 ms (average p value = 0.01). Brainstem responses also significantly increased, within a window of 13 to 42 ms (average p value = 0.02) and 15 to 43 ms (average p value = 0.02) for responses elicited by even and odd beats respectively (Fig. 5).

In contrast, for both cortical and brainstem responses, there were no differences between the neural responses to the even and odd beats (all p values = 1), neither in the auditory, nor in the tapping condition. Hence, while the amplitude of these responses was enhanced in the tapping condition as compared to the auditory condition, this enhancement did not systematically follow the movement cycles over the sequence.



**Figure 5.** Time course of the cortical and brainstem SS-EPs (group-level average). The 416-ms chunks (corresponding to the window between each beat onset) were separated into odd and even beats (i.e. beats with and without synchronized tapping respectively). For both the cortical responses (obtained by filtering the EEG between 0.3 and 30 Hz) and the brainstem responses (obtained by filtering the EEG between 190 and 610 Hz, then by extracting the envelope of the filtered signal using a Hilbert transform), no significant differences were found between odd and even beats (solid and dashed lines). In contrast, comparison of the time courses across conditions yielded significant differences in similar time windows between odd and even beats (grey boxes).

#### 4. DISCUSSION

Participants listened to a chord made of three partial tones whose envelope was modulated in amplitude to induce a steady periodic beat. Concurrent auditory brainstem and cortical responses to the sound were isolated in the EEG signal based on their frequencies using a frequency-tagging approach. When the participants tapped in synchrony on the beat, the cortical responses at beat frequency and harmonics were larger than during passive listening. Importantly, brainstem responses at the frequencies of the partial and sideband components were also enhanced. However, this gain in amplitude during tapping was not uniformly distributed over the different components of the brainstem response. Rather, we observed a sharpened amplification for the sideband components as compared to the partials, located in particular at the lower-pitched tone of the chord. To our knowledge, these results constitute the first evidence of a gain in amplitude of auditory brainstem responses when synchronizing movements to a sound. Moreover, they are compatible with the view that sensorimotor synchronization implies an online modification of the encoding of the sound not merely in the cortex but also in the brainstem, in particular for the auditory features that are relevant to the sensorimotor coupling behavior.

##### 4.1. Enhanced brainstem encoding: preferential encoding of the envelope

Our observations provide evidence for an enhancement of the brainstem encoding of sound during sensorimotor synchronization. Specifically, the sideband components of the brainstem response were selectively enhanced as compared to the partial components. This enhancement could be interpreted as a top-down modulation from cortical areas descending to subcortical nuclei through corticofugal tracks (Bajo and King, 2013). In line with this interpretation, it has recently been proposed to consider the auditory system as a series of dynamic loops in which the neural encoding of sounds in the subcortical nuclei is

continually updated by changes in activity at higher levels, thus shaping in turn the cortical processing of auditory information (Bajo and King, 2013). A mechanistic explanation for such sensitivity in the brainstem encoding of sound could be that in the current study efferent corticofugal projections transmitted the gain in amplitude of the cortical response frequency-locked to the beat back to the inferior colliculus. This interpretation is supported by previous studies having demonstrated that the brainstem encoding of sounds as gathered from the human scalp is not invariant. Rather, there is increasing evidence for task-dependent modulations of auditory brainstem encoding explained by the existence of efferent corticofugal projections (Musacchia et al., 2007; Chandrasekaran et al., 2009; Lehmann and Schönwiesner, 2014).

What is the nature of the sensory gain observed for both cortical and subcortical responses to sounds in the tapping condition remains an open question. For instance, it could be hypothesized that explicit motor or tactile factors were responsible for this enhancement, via corollary discharges due to movement command or somatosensory feedback transmitted to the brainstem through the corticofugal tracks. According to this interpretation, the time course of the sensory gain would be expected to relate to the time course of the movement. However, frequency and time domain analyses of the responses to the sound envelope revealed no significant modulation corresponding to the movement.

Alternatively, the sensory gain observed in the current study could also reflect a general modulation of the state of arousal caused by the concurrent production of body movement, resulting in a larger response to the sound input by the auditory neurons throughout the ascending auditory pathway (see e.g. McGinley et al., 2015, for a recent evidence of optimal arousal states for sensory signal detection in non-human animals). According to this hypothesis, the cortical and subcortical sensory gain in response to the sound input would

be expected to arise regardless of the fact that the movement is synchronized or not to the sound input. Moreover, the sensory gain would be expected to exhibit a uniform distribution over the different components of the sound input (i.e. no selective gain of the sound envelope vs. sound partials).

Rather, the enhanced brainstem encoding of the envelope observed in the tapping condition could reflect a general mechanism of enhanced encoding of behavior-relevant components of the incoming sensory stream (corresponding here to the sound envelope to which the tapping were synchronized). Such mechanism of active sensing has been proposed to rely on a feedforward high-accuracy pathway (e.g. via connections with the cerebellum) and feedback modulations through descending corticofugal pathways (see e.g. Schwartze and Kotz, 2013). Along this line, brainstem responses to sounds have been shown to be modulated by selective attention (Lehmann and Schönwiesner, 2014; Chandrasekaran et al., 2009). In the current study, the sensorimotor synchronization component of the task might thus constitute a means to drive a selective sharpening of the behavior-relevant feature of the auditory input through a process of selective dynamic attention, as proposed by recent models of active sensing (see e.g. Morillon et al., 2014). This interpretation fits well with our current observations of a selective enhancement of the encoding of the sound envelope, particularly at the lower-tone. Future research should clarify the selectivity of the sensory gain by comparing experimental conditions in which the participant is asked to focus on distinct behavior-relevant components (sound envelope vs. sound pitch for instance). Moreover, the current study provides a new valid experimental design to investigate the model in cerebellar patients, with the hypothesis of an impaired selective gain of the sidebands as compared to control participants in the tapping condition. Finally, future studies are needed to compare extensively with the same method different kinds of body

movement vs other kinds of dynamic sensory inputs for instance in their ability to drive a selective shaping of the auditory input in real time.

#### **4.2. Preferential encoding of low-pitched tones: the timing cues hypothesis**

In the current study, the gain in amplitude of the sound encoding during movement was particularly salient in the lower-pitched tone of the chord. This observation could be interpreted as a direct cortical effect, if we consider the responses frequency-locked at 200 Hz and sidebands as a mix of cortical and brainstem responses. Indeed, it could be hypothesized that EEG responses frequency-locked to sounds could actually consist in a mixture of activities originating not merely from brainstem auditory nuclei but also from cortical neurons, with a relative weight moving towards a smaller influence of the cortical sources in favor of brainstem auditory sources at higher frequencies. However, this is unlikely to be the case in our current study, based on the analysis of the phase of the responses. Indeed, the relative phase of the 2.4 Hz modulation measured in the 200-600 Hz frequency band (mostly driven by the 200 Hz response because of its higher amplitude than the 400 and 600 Hz responses) was significantly different from the relative phase of the cortical activity at 2.4 Hz. Hence, this result suggests that the activity measured at 200 Hz and sidebands originates mainly from brainstem to the detriment of cortical sources.

Alternatively, our current finding of a gain in amplitude of the brainstem sideband components particularly located at the lower-pitched tone of the chord when synchronizing to the sound could be explained by a low-tone/timing connection hypothesis. Indeed, there is increasing evidence that bass frequencies have a pronounced effect on body movement (Hove and Keller, 2014). For instance, low-pitched frequencies were associated with increased temporal regularity of movement in motion-capture studies of moving to music (Van Dyck et al., 2013). When participants synchronized tapping movements with chords

containing small onset asynchronies, tap timing was more strongly influenced by the lower-pitched tone (Hove et al., 2007). Moreover, the tendency to move along with the lower-pitched tones is reflected by the musical convention that bass-ranged instruments often lay down the rhythmic foundation and provide the pulse on metrically strong beats (Large, 2002). Recently, an EEG study found that cortical responses to timing deviants randomly inserted in a sequence of repeated chords were larger for the lower compared to the higher tone of the chord, thus suggesting a more robust temporal encoding of the lower-pitched stream (Hove et al., 2014). A possible mechanistic explanation for this greater influence of the lower tone on auditory–motor synchronization was recently proposed based on a biologically plausible model of the auditory periphery (Zilany et al., 2009). The modeling revealed that superior time encoding for the lower-pitched tones already arises at peripheral level, namely in the cochlea of the inner ear due to specificities of its functional organization (Hove et al., 2014). Together with these studies, our current finding of a gain in amplitude of the brainstem sideband components particularly located at the lower-pitched tone of the chord when synchronizing to the sound provides support to this low-tone/timing connection hypothesis.

#### **4.3. Enhanced cortical encoding: motor overlap confound vs. auditory-motor interaction**

An enhancement of the cortical responses frequency-locked to the beat envelope was observed during sensorimotor synchronization as compared to the auditory condition, thus confirming previous EEG findings (Nozaradan et al., 2015). This enhancement could be explained by the motor cortical responses overlapping in part with the auditory activity frequency-locked to the beat at the electrodes of interest. Indeed, a substantial peak of electromyographic activity was measured at the beat frequency (2.4 Hz) in addition to the expected peak measured at 1.2 Hz corresponding to tapping movement on alternating

beats. In turn, a significant activity was observed at 1.2 Hz and harmonics on the electrodes of interest, which could reflect this motor overlap confound. In fact, this confound cannot be totally excluded, due to the uncertainty inherent to scalp EEG on the possible sources of the activities gathered from the head surface (Nozaradan et al., 2015). However, several arguments make the motor confound hypothesis unlikely to explain the gain in amplitude of the cortical beat SS-EPs when tapping. For example, the amplitude of the activity at 1.2 Hz and harmonics was found to be much smaller than the amplitude elicited by the sound envelope at 2.4 Hz. Moreover, the measures were done on electrodes located on the median line of the scalp, i.e. on electrodes that do not correspond to the left-lateralized electrodes typically used to capture right hand motor-evoked potentials (Bourguignon et al., 2015). Furthermore, the signals were aligned to the sound onset rather than to each tapping, such as to optimize the measure of cortical responses specific to the sound to the detriment of motor responses. Also, in the case of a significant motor confound, a significant difference across conditions could be expected in the relative phases of the cortical beat SS-EPs, due to the different latencies and variability of motor-evoked potentials. However, no significant difference in the phase of the cortical SS-EP at 2.4 Hz was observed across the two conditions. Finally, source analysis of cortical beat SS-EPs conducted with similar experimental parameters than in the current study has provided evidence for sources compatible with auditory cortical areas rather than movement-related areas for these cortical beat SS-EPs (Nozaradan et al., 2015). All these points argue in favor of the view that the cortical gain *and* the small activity at 1.2 Hz and harmonics observed in the tapping condition are actual effects of the task (auditory-motor synchronization process) rather than merely an effect of motor overlap confound.

Alternatively, the observed increase in frequency-locking to the sound envelope during movements may reflect a mechanism of temporal selective attention (Calderone et al., 2014; Lakatos et al., 2008). Indeed, a dynamic attention process, possibly supported by an underlying selective neural entrainment to the behavior-relevant stream, may enable listeners to better predict when sound features are likely to occur, thus facilitating overt synchronization (Tierney and Kraus, 2014; Nozaradan et al., 2016). Hence, if a movement synchronization to a repeated beat involves a process of selective frequency-locking of neural populations to the beat time-course (Large, 2008), this would yield enhanced SS-EP amplitudes at beat frequency (Chemin et al., 2014; Nozaradan et al., 2015). According to this view, it has recently been shown that during sensorimotor synchronization to a steady beat the tapping behavior produces cyclic fluctuations in sensory gain that are locked to individual movements (Merchant et al., 2015). This evidence suggests that signals aligned to rhythmic motor execution may tune representations of incoming sensory information, thus explaining the additional 1.2 Hz activity observed here during movement.

#### **4.4. Conclusion: relationship between sensorimotor synchronization and sensory gain**

The current results support the view that, although the human auditory pathway shows a remarkable ability to synchronize to the incoming acoustic input over time, these synchronous neural responses do not merely constitute an invariant encoding of the input. Rather, the system transforms the dynamic input already at subcortical levels by amplifying features that are relevant for perception and behavior. This dynamic shaping would thus allow to selectively lock individuals' mind and body to incoming rhythmic sensory streams. While the frequency-tagging method appears to be revealing about the neural encoding of these behavior-relevant sensory features, the functional relevance of this enhanced encoding during the actual behavior remains to be addressed directly. For instance, further

research using the same approach but tracking movement kinematics more extensively (e.g. via measures of force, trajectory, accuracy and precision) across a sample of individuals with large variance in sensorimotor synchronization performance (e.g. musicians and nonmusicians) could address this question more adequately, by testing the correlation between the gain in sensory encoding and individual differences in tapping performance. Hence, our study providing first evidence of enhanced cortical and subcortical encoding of sounds during movement opens to future studies investigating the ability of dynamic inputs (including movement) to drive a selective shaping of sound encoding in real time.

Acknowledgements: S.N. is supported by the Fund for Scientific Research of the French-speaking community of Belgium (F.R.S.-FNRS). MS was supported by a career grant from the Fonds de recherche du Québec – Santé (FRQS).

## REFERENCES

- Bajo VM, King AJ (2013) Cortical modulation of auditory processing in the midbrain. *Front Neural Circuits* 6:114.
- Bengtsson SL, Ullén F, Ehrsson HH, Hashimoto T, Kito T, Naito E, Forssberg H, Sadato N (2009) Listening to rhythms activates motor and premotor cortices. *Cortex* 45(1):62-71.
- Berens P (2009) CircStat: A Matlab Toolbox for Circular Statistics. *Journal of Statistical Software* 31:10.
- Bidelman, GM (2013) The role of the auditory brainstem in processing musically relevant pitch. *Front Psychol* 4:264.
- Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J Neurosci* 27(35):9252-61.
- Bijsterbosch JD, Lee K, Hunter MD, Tsoi DT, Lankappa S, Wilkinson ID, Barker AT, Woodruff PWR (2010) The role of the cerebellum in sub- and supraliminal error correction during sensorimotor synchronization: evidence from fMRI and TMS. *J Cogn Neurosci* 23:1100-1112.
- Bourguignon M, Piitulainen H, De Tiège X, Jousmäki V, Hari R (2015) Corticokinematic coherence mainly reflects movement-induced proprioceptive feedback. *Neuroimage* 106:382-90.
- Calderone DJ, Lakatos P, Butler PD, Castellanos FX (2014) Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn Sci* 18(6):300-9.
- Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N (2009) Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* 64(3):311-9.
- Chemin B, Mouraux A, Nozaradan S (2014) Body movement selectively shapes the neural representation of musical rhythm. *Psychological Science* pii: 0956797614551161.
- Diamond IT, Jones EG, Powell TP (1969) The projection of the auditory cortex upon the diencephalon and brain stem in the cat. *Brain Res* 15(2):305-40.
- Doumas M, Praamstra P, Wing AM (2005) Low frequency rTMS effects on sensorimotor synchronization. *Exp Brain Res* 167:238-45.
- Drake C, Botte MC (1993) Tempo sensitivity in auditory sequences: Evidence for a multiple look model. *Percept Psychophys* 54:277-286.
- Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A* 78(4):2643-7.
- Galbraith GC, Bhuta SM, Choate AK, Kitahara JM, Mullen TA Jr (1998) Brain stem frequency-following response to dichotic vowels during attention. *Neuroreport* 9(8):1889-93.

- Gerloff C, Uenishi N, Nagamine T, Kunieda T, Hallett M, Shibasaki H (1998) Cortical activation during fast repetitive finger movements in humans: steady-state movement-related magnetic fields and their cortical generators. *Electroencephalogr Clin Neurophysiol* 109:444-53.
- Greenberg S, Marsh JT, Brown WS, et al. (1987) Neural temporal coding of low pitch. I. Human frequency following responses to complex tones. *Hear Res* 25:91–114.
- Groppe DM, Urbach TP, Kutas M (2011) Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48(12):1711-25.
- Hattori Y, Tomonaga M, Matsuzawa T (2015) Distractor Effect of Auditory Rhythms on Self-Paced Tapping in Chimpanzees and Humans. *PLoS One* 10(7):e0130682.
- Hove MJ, Balasubramaniam R, Keller PE (2014) The time course of phase correction: a kinematic investigation of motor adjustment to timing perturbations during sensorimotor synchronization. *J Exp Psychol Hum Percept Perform* 40(6):2243-51.
- Hove MJ, Keller PE, Krumhansl CL (2007) Sensorimotor synchronization with chords containing tone-onset asynchronies. *Attention Percept Psychophysiol* 69:699-708.
- Hove MJ, Marie C, Bruce IC, Trainor LJ (2014) Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proc Natl Acad Sci U S A* 111:10383-8.
- Jäncke L, Loose R, Lutz K, Specht K, Shah NJ (2000) Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cogn Brain Res* 10(1-2):51-66.
- Kay SM (1988) *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, USA.
- Krishnan, A (2007) Frequency-following response. *Auditory evoked potentials: Basic principles and clinical application*. In: Burkard RF, Eggermont JJ, Don M (eds) Lippincott Williams and
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320(5872):110-3.
- Large E (2002) Perceiving temporal regularity in music. *Cogn Sci* 26:1-37.
- Large EW (2008) Resonating to musical rhythm: Theory and experiment. In Simon Grondin (Ed.) *The Psychology of Time*. Emerald, West Yorkshire, USA.
- Lehmann A, Schönwiesner M (2014) Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PLoS One* 9(1):e85442.
- Madison G and Sioros G (2014) What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Front Psychol* 5:894. doi: 10.3389/fpsyg.2014.00894.

- Maris E and Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177-90.
- McGinley MJ, David SV, McCormick DA (2015) Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* 87(1):179-92.
- Meigen T and Bach M (1999) On the statistical significance of electrophysiological steady-state responses. *Doc Ophthalmol* 98(3):207-32.
- Merchant H, Pérez O, Bartolo R, Méndez JC, Mendoza G, et al (2015) Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *Eur J Neurosci* 41:586-602.
- Morillon B, Schroeder CE, Wyart V (2014) Motor contributions to the temporal precision of auditory attention. *Nat Commun* 5:5255.
- Morillon B, Schroeder CE, Wyart V (2014) Motor contributions to the temporal precision of auditory attention. *Nat Commun* 5:5255.
- Mouraux A, Iannetti GD, Colon E, Nozaradan S, Legrain V, Plagkhi L (2011) Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. *J Neurosci* 31:6079-87.
- Mulders WH, Robertson D (2000) Evidence for direct cortical innervation of medial olivocochlear neurones in rats. *Hear Res* 144(1-2):65-72.
- Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci U S A* 104(40):15894-8.
- Néda Z, Ravasz E, Brechet Y, Vicsek T, Barabási AL (2000) The sound of many hands clapping. *Nature* 403(6772):849-50.
- Norcia AM, Appelbaum LG, Ales JM, Cottureau BR, Rossion B (2015) The steady-state visual evoked potential in vision research: a review. *J Vis* 15(6):4 doi: 10.1167/15.6.4.
- Nozaradan S (2014) Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Phil Trans B* doi:10.1098/rstb.2013.0393.
- Nozaradan S, Peretz I, Keller PE (2016) Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. *Nature Scientific Reports*. In press.
- Nozaradan S, Peretz I, Missal M, Mouraux A (2011) Tagging the neuronal entrainment to beat and meter. *J Neurosci* 31:10234-40.
- Nozaradan S, Peretz I, Mouraux A (2012) Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J Neurosci* 32(49):17572-81.
- Nozaradan S, Peretz I, Mouraux A (2012) Steady-state evoked potentials as an index of multisensory temporal binding. *NeuroImage* 60:21-8.

- Nozaradan S, Zerouali Y, Peretz I, Mouraux A (2015) Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cereb Cortex* 25:736-47.
- Patel AD and Iversen JR (2014) The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Psychology* 8:56 doi: 10.3389/fpsyg.2014.00057.
- Penhune VB, Zattore RJ, Evans AC (1998) Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J Cogn Neurosci* 10(6):752-65.
- Picton TW, Skinner CR, Champagne SC, Kellett AJ, Maiste AC (1987) Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *J Acoust Soc Am* 82(1):165-78.
- Rees A, Green GG, Kay RH (1986) Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hear Res* 23(2):123-33.
- Regan DM (1989) *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Eds Regan DM, Elsevier, USA.
- Repp BH (2005) Sensorimotor synchronization: A review of the tapping literature. *Psychon B Rev* 12:969-992.
- Repp BH, Su YH (2006-2012) Sensorimotor synchronization: a review of recent research. *Psychon Bull Rev* 20(3):403-52.
- Rinne T, Balk MH, Koistinen S, Autti T, Alho K, Sams M (2008) Auditory selective attention modulates activation of human inferior colliculus. *J Neurophysiol* 100(6):3323-7.
- Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *J Acoustical Soc of America* 108:679-691.
- Ross B, Picton TW, Herdman AT, Pantev C (2004) The effect of attention on the auditory steady-state response. *Neurol Clin Neurophysiol* 2004:22.
- Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of Active Sensing and perceptual selection. *Curr Opin Neurobiol* 20(2):172-6.
- Schröger E, Marzecová A, SanMiguel I (2015) Attention and prediction in human audition: a lesson from cognitive psychophysiology. *Eur J Neurosci* 41(5):641-64.
- Schubotz RI (2007) Prediction of external events with our motor system: towards a new framework. *Trends Cogn Sci* 11(5):211-8.
- Schwartz M, Korz SA (2013) A dual-pathway neural architecture for specific temporal prediction. *Neurosci Biobehav Rev* 37:2587-96.

Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31(3):302-24.

Skoe E, Kraus N (2010) Hearing it again and again: on-line subcortical plasticity in humans. *PLoS One*. 5(10):e13645.

Slee SJ, David SV (2015) Rapid Task-Related Plasticity of Spectrotemporal Receptive Fields in the Auditory Midbrain. *J Neurosci* 35(38):13090-102.

Tierney A, Kraus N (2013) The ability to move to a beat is linked to the consistency of neural responses to sound. *J Neurosci* 33(38):1481-8.

Tierney A, Kraus N (2014) Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Front Hum Neurosci* 8:949.

Timm J, SanMiguel I, Keil J, Schröger E, Schönwiesner M (2014) Motor intention determines sensory attenuation of brain responses to self-initiated sounds. *J Cogn Neurosci* 26(7):1481-9.

Ullén F, Forsman L, Blom O, Karabanov A, Madison G (2008) Intelligence and variability in a simple timing task share neural substrates in the prefrontal white matter. *J Neurosci* 28(16):4238-43.

van den Broeke EN, Mouraux A, Groneberg AH, Pfau DB, Treede RD, Klein T (2015) Characterizing pinprick-evoked brain potentials before and after experimentally induced secondary hyperalgesia. *J Neurophysiol* 114(5):2672-81.

Van Dyck ED, Moelants M, Demey M et al (2013) The impact of the bass drum on human dance movement. *Music Percept* 30:349-359.

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ (2012) Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *J Neurophysiol* 107(8):2033-41.

Weedman DL, Ryugo DK (1996) Projections from auditory cortex to the cochlear nucleus in rats: synapses on granule cell dendrites. *J Comp Neurol* 371(2):311-24.

Wilkins, Philadelphia p. 313-335.

Winer JA (2006) Decoding the auditory corticofugal systems. *Hear Res* 212(1-2):1-8.

Winer JA, Larue DT, Diehl JJ, Hefti BJ (1998) Auditory cortical projections to the cat inferior colliculus. *J Comp Neurol* 400(2):147-74.

Winer JA, Lee CC (2007) The distributed auditory cortex. *Hear Res* 229(1-2):3-13.

Wong WYS and Stapells DR (2004) Brain stem and cortical mechanisms underlying the binaural masking level difference in humans: and auditory steady-state response study. *Ear & Hearing* 25:57-67.

Zagha E, Casale AE, Sachdev RN, McGinley MJ, McCormick DA (2013) Motor cortex feedback influences sensory processing by modulating network state. *Neuron* 79(3):567-78.

Zilany MSA, Bruce IC, Nelson PC, Carney LH (2009) A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *J Acoust Soc* 126:2390-2412.

ACCEPTED MANUSCRIPT

**Highlights**

- enhanced cortical and subcortical encoding of sound during synchronized movement
- selective boost of behavior-relevant sound features in the brainstem when moving
- the brainstem encoding of bass tones is selectively boosted during sync movement

ACCEPTED MANUSCRIPT