

Neural Entrainment to the Beat: The “Missing-Pulse” Phenomenon

Idan Tal,¹ Edward W. Large,² Eshed Rabinovitch,¹ Yi Wei,² Charles E. Schroeder,^{3,4} David Poeppel,^{5,6} and Elana Zion Golumbic¹

¹Gonda Multidisciplinary Brain Research Center, Bar Ilan University, Ramat Gan 5290002, Israel, ²Department of Psychology, University of Connecticut Storrs, Connecticut 06269, ³Department of Psychiatry, Columbia University Medical Center, New York, New York 10032, ⁴Translational Cognitive Neuroscience Program, Nathan Kline Institute, Orangeburg, New York 10962, ⁵Department of Psychology, New York University, New York, New York 10003, and ⁶Neuroscience Department, Max-Planck Institute for Empirical Aesthetics, Frankfurt, Germany 60322

Most humans have a near-automatic inclination to tap, clap, or move to the beat of music. The capacity to extract a periodic beat from a complex musical segment is remarkable, as it requires abstraction from the temporal structure of the stimulus. It has been suggested that nonlinear interactions in neural networks result in cortical oscillations at the beat frequency, and that such entrained oscillations give rise to the percept of a beat or a pulse. Here we tested this neural resonance theory using MEG recordings as female and male individuals listened to 30 s sequences of complex syncopated drumbeats designed so that they contain no net energy at the pulse frequency when measured using linear analysis. We analyzed the spectrum of the neural activity while listening and compared it to the modulation spectrum of the stimuli. We found enhanced neural response in the auditory cortex at the pulse frequency. We also showed phase locking at the times of the missing pulse, even though the pulse was absent from the stimulus itself. Moreover, the strength of this pulse response correlated with individuals' speed in finding the pulse of these stimuli, as tested in a follow-up session. These findings demonstrate that neural activity at the pulse frequency in the auditory cortex is internally generated rather than stimulus-driven. The current results are both consistent with neural resonance theory and with models based on nonlinear response of the brain to rhythmic stimuli. The results thus help narrow the search for valid models of beat perception.

Key words: auditory rhythm; MEG; neural resonance theory; oscillations; pulse

Significance Statement

Humans perceive music as having a regular pulse marking equally spaced points in time, within which musical notes are temporally organized. Neural resonance theory (NRT) provides a theoretical model explaining how an internal periodic representation of a pulse may emerge through nonlinear coupling between oscillating neural systems. After testing key falsifiable predictions of NRT using MEG recordings, we demonstrate the emergence of neural oscillations at the pulse frequency, which can be related to pulse perception. These findings rule out alternative explanations for neural entrainment and provide evidence linking neural synchronization to the perception of pulse, a widely debated topic in recent years.

Introduction

Music often contains complex temporal patterns, with no simple regularity predicting when a particular note will occur. Nonethe-

less, listeners typically perceive music as having a regular pulse, or beat, marking equally spaced points in time within which sequences of musical notes are organized. Moreover, an almost automatic response is to move to the beat of the music (Nettl, 2000; Michaelis et al., 2014). The process through which a percept of pulse emerges is referred to as induction, and most listeners are quite accurate in identifying and tapping to a pulse, even for highly nonisochronous rhythms (Large et al., 2002, 2015; Patel et al., 2005).

The neural mechanisms underlying pulse induction are not fully understood, but it is clear that the phenomenon involves

Received Aug. 7, 2016; revised May 9, 2017; accepted May 16, 2017.

Author contributions: I.T., E.W.L., C.E.S., D.P., and E.Z.G. designed research; I.T., E.R., and E.Z.G. performed research; I.T., E.W.L., E.R., Y.W., and E.Z.G. analyzed data; I.T., E.W.L., C.E.S., D.P., and E.Z.G. wrote the paper.

This work was supported by the I-CORE (Israeli Centers of Research Excellence) Program of the Planning and Budgeting Committee and the Israel Science Foundation (Grant 51/11; E.Z.G.), the FP7 Marie Curie Career Integration Grant (E.Z.G.), the Bi-National Science Foundation (Grant 2013232; E.Z.G. and D.P.), and the National Institutes of Health (Grant R01DC05660; D.P.; Grant MH103814; C.E.S.).

The authors declare no competing financial interests.

Correspondence should be addressed to Elana Zion Golumbic, Gonda Multidisciplinary Brain Research Center, Bar Ilan University, Ramat Gan 5290002, Israel. E-mail: elana.zion-golumbic@biu.ac.il.

DOI:10.1523/JNEUROSCI.2500-16.2017

Copyright © 2017 the authors 0270-6474/17/376331-11\$15.00/0

complex dynamics in multiple brain regions, including the superior temporal gyrus (STG), the supplementary motor area (SMA), the frontal cortex, the basal ganglia, and the cerebellum (Grahn and Brett, 2007; Bengtsson et al., 2009; Chen et al., 2009; Chapin et al., 2010; Merchant et al., 2015). Large and colleagues recently proposed a computational model, neural resonance theory (NRT), for explaining how nonperiodic stimuli interact with endogenous brain rhythms to induce the percept of a periodic pulse (Snyder and Large, 2005; Large, 2008; Large and Snyder, 2009). According to NRT, the pulse emerges through nonlinear coupling between two oscillatory networks—one representing the physical properties of the stimulus (labeled “sensory”), and a second (labeled “motor”) that integrates inputs from the sensory system. The nonlinear interactions give rise to oscillatory activity not only at the frequencies present in the stimulus, but also at more complex combinations, including the pulse frequency (and its subharmonics).

A critical test for empirical evaluation of NRT is to use syncopated rhythms in which the pulse frequency is not physically present in the stimulus envelope. It has been shown that syncopated rhythms containing no energy at the pulse frequency in their broadband modulation spectrum (Velasco and Large, 2011) can nonetheless induce the percept of a pulse (Chapin et al., 2010; Large et al., 2015). Such rhythms are created using phase cancellation—half of the notes occur in-phase with the intended pulse, the other half occur in anti-phase. Here we tested whether the neural response to such stimuli contains energy at the pulse frequency, despite its physical absence in the acoustics. If this were the case, it would support the notion that pulse induction emerges through nonlinear interaction between the stimulus and endogenous neural dynamics, and would rule out simpler explanations of apparent entrainment stemming either from transient responses to individual rhythmic events (Tremblay et al., 2004; Sussman et al., 2008) or from linear frequency analysis of the stimuli (Todd, 1999; Tomic and Janata, 2008). Previous studies have demonstrated that the steady-state EEG response to syncopated stimuli shows enhanced power at the pulse frequency, at magnitudes exceeding what is expected from the acoustic structure of the stimulus alone, which is in line with NRT predictions (Nozaradan et al., 2011, 2012, 2016). However, in most of those studies, the pulse frequency was either present in the stimulus envelope, or was induced through instruction, making it difficult to rule out the possibility that these effects reflect selective amplification of neural responses to on-beat sounds (Tierney and Kraus, 2014; Bouwer and Honing, 2015).

In the current study, using MEG recordings while individuals listened to “pulseless” syncopated rhythms, we directly tested the hypotheses posed by NRT that oscillations at the pulse frequency would emerge independent of stimulus acoustics. We were also interested in testing whether the emergence of a neural representation for the pulse was linked to the behavioral perception of a pulse, since previous studies have suggested that pulse induction for syncopated rhythms is not “automatic,” but requires attention (Chapin et al., 2010) and can be modified by instruction (Large et al., 2002; Nozaradan et al., 2011).

Materials and Methods

Experiment 1

Participants. Eleven subjects (six males; mean age, 30 years; SD, 3.7 years) received payment to participate in the study and signed a consent form approved by the New York University Institutional Review Board ethics committee. Participants did not suffer from any neurological disorders and had normal hearing.

MEG recordings. MEG data were collected on a 157-channel whole-head MEG system [5 cm baseline axial gradiometer superconducting

quantum interference device (SQUID)-based sensors (Kanazawa Institute of Technology)] in a magnetically shielded room (Vakuum-schmelze). Data were sampled at 1000 Hz, with a notch filter at 60 Hz, and an on-line recording 200 Hz low-pass filter. Using 3D digitizer software (Source Signal Imaging) and digitizing hardware (Polhemus), each participant’s head position was assessed via five coils attached to anatomical landmarks both before and after the experiment to ensure that head movement was minimal. The auditory signals were presented through in-ear earphones (ER3-A, Etymotic Research) at comfortable, conversational levels (~72 dB sound pressure level). Stimulus delivery and triggering was controlled by the Presentation program (Neurobehavioral Systems).

Stimuli. Stimuli consisted of sequences of drumbeats, with a rise time of 0.5 ms (calculated as the time it took the acoustic signal to cross from 10 to 90% reference levels) and pulse width of 20 ms. Four different sequences were used, differing in their temporal pattern, as illustrated in Figure 1*a*. The isochronous sequence (ISO) contained eight drumbeats presented with a constant interonset interval (IOI) of 0.5 s. Two syncopated patterns that contained no energy at the pulse frequency when measured using linear analysis and are thus referred to as missing pulse (MP1 and MP2) were selected from the stimuli used by Chapin et al. (2010). These complex rhythmic patterns contained eight drumbeats and were constructed so that half of the drumbeats occurred on “strong” beats and half on “weak” beats of a 2 Hz rhythm. These highly syncopated patterns are expected to give rise to the perception of a 1 or 2 Hz pulse, yet the Fourier spectrum of the stimuli contain no energy at these frequencies. The random condition (RAND) consisted of eight drumbeats placed in a temporally random sequence that did not induce any perception of a pulse. All sequences had an equal length of 4 s. We analyzed the modulation spectrum of the stimulus acoustics by computing the broadband amplitude envelope using the absolute value of the Hilbert transform. Next, we computed the spectrum of the envelope using FFT on two cycles of each sequence (8 s), yielding a frequency resolution of 0.125 Hz. We used the broadband envelope to assess the modulation spectrum. In the general case, the cortical inputs relevant to beat perception would more realistically be described as event onsets in multiple cochlear frequency channels (Langner, 1992). However, since the percussive sounds used here as stimuli are brief and broadband, the two input representations yield nearly identical frequency-domain descriptions. This simpler representation enabled us to apply a similar analysis to the stimulus and the neural signal.

As seen in Figure 2 (top), neither syncopated stimuli contained a peak at 1 or 2 Hz, frequencies perceived as the pulse of these rhythms. One should note that although both syncopated rhythms contained instances where the interval between sounds is precisely 500 ms (two such intervals in MP1; three such intervals in MP2), these intervals were never consecutive. Consequently, due to the uncertainty principle, there is no principled way for a linear system to produce a 2 Hz response, even if a shorter time window is used to calculate the modulation spectrum.

Experimental procedure. The time course of a sample trial is illustrated in Figure 1*b*. Participants listened to eight consecutive repetitions of one of the patterns (total 32 s). Following this passive listening period, the participants heard an additional test item consisting of one repetition of the same pattern but possibly with a different tempo (same tempo, 10% faster, or 10% slower, equiprobable). Participants were then asked to judge whether the tempo of the test item was faster, slower, or the same as the original pattern. This was done to keep the subjects engaged while passively listening to the long auditory stimulus and staying attentive to its temporal properties without explicitly instructing them to extract a pulse. Trials with the four different temporal conditions (ISO, MP1, MP2, RAND) were presented in pseudorandom order and each condition was repeated 10–15 times throughout the experiment.

MEG preprocessing and analysis. All analyses were performed using Matlab (Mathworks), as well as the Fieldtrip toolbox (Oostenveld et al., 2011). The data were noise-reduced off-line using a time-shift principled component analysis (de Cheveigné and Simon, 2007). Heartbeat, eye movement, and blink artifacts were removed from the data using an Independent Component Analysis algorithm from the Fieldtrip toolbox.

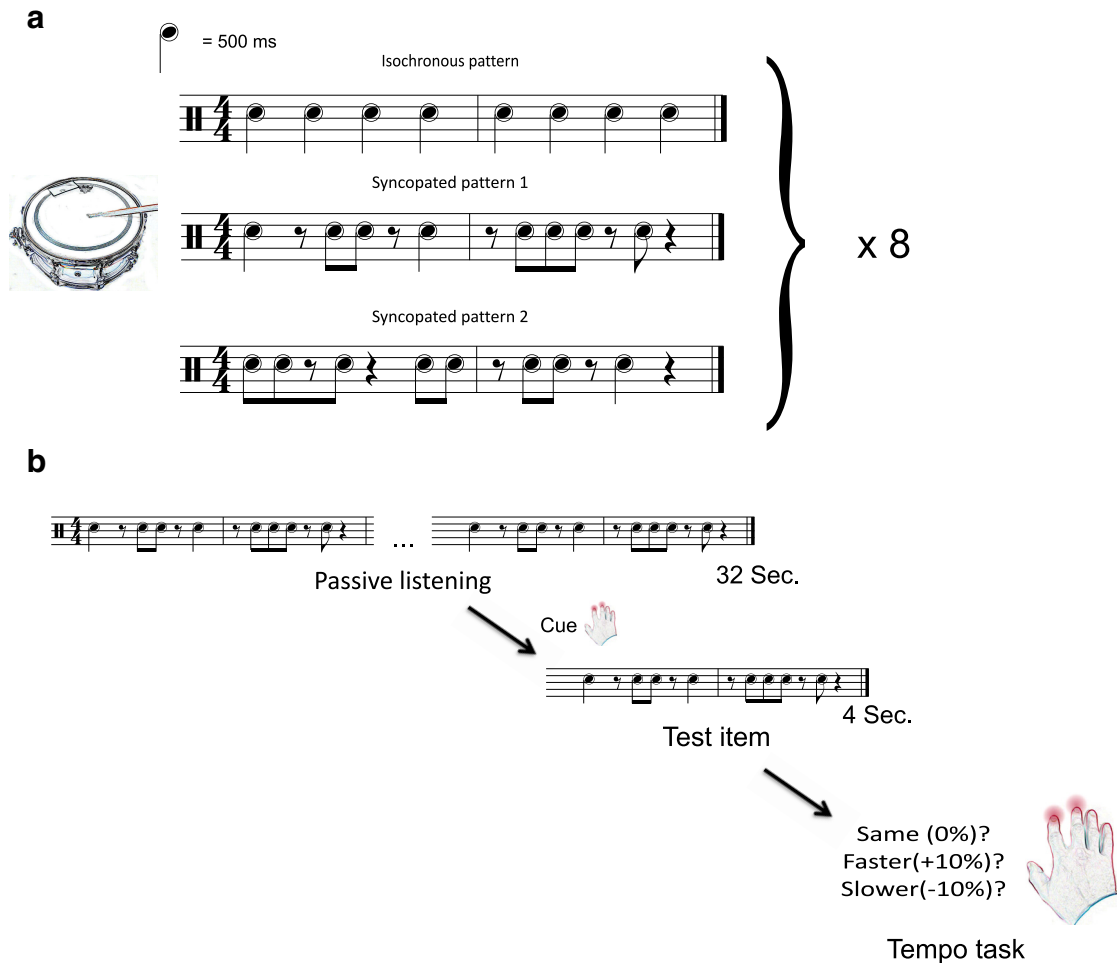


Figure 1. Stimulus and paradigm. **a**, Notation of the auditory stimuli for the isochronous pattern (top), the syncopated pattern MP1 (middle), and the syncopated pattern MP2 (bottom) stimuli. **b**, Time course of a sample trial. In the passive listening period, the basic pattern (2 musical bars) was repeated eight times. Then, following a cue, a test item was played consisting of the same basic pattern but the tempo could be the same, 10% faster, or 10% slower. Participant performed a tempo-judgment task on the test item.

Additional visual inspection of the data was performed to remove data segments with large muscle or external artifacts.

Spectral analysis. MEG signals were segmented into 32 s epochs corresponding to the “passive listening” periods, and divided into four conditions. The first 1 s of data was removed from all epochs, to avoid onset effects. Epochs were then averaged for each condition, and the spectrum of the averaged data (zero padded) was calculated using Fourier transform at each of the 157 sensors, separately for each participant. Visual inspection revealed a peak in the MEG spectrum at 2 Hz that did not exist in the stimulus spectrum and matches the preferred intrinsic pulse tempo for these syncopated stimuli (Moelants, 2002; Honing, 2013; Large et al., 2015).

For statistical analysis of the pulse-frequency peaks observed in the MEG spectrum, we averaged the MEG responses for each subject over all sensors. We then calculated the nonparametric Wilcoxon signed-rank test between the 2 Hz power and the mean power at frequencies around it (1.875 and 2.125 Hz). We specifically targeted this frequency a priori (as the stimuli were constructed to induce perception of a pulse at this specific rate), and the RAND condition served as a control to test that the observed effects were not biased due to this a priori frequency selection.

Phase-locking analysis. The stimuli were constructed so that half of the drumbeats occurred on strong-beat positions, whereas in half of the strong-beat positions no stimuli occurred. This allowed us to look directly at the neural activity at positions uncontaminated by sensory evoked potentials, in addition to analyzing the spectrum of the neural response to the entire continuous stimuli. To this end, single-trial data were filtered between 0.5 and 3 Hz and the instantaneous phase at each

time point was extracted using the Hilbert transform (implemented in Matlab). For each of the syncopated conditions, we averaged the phase values for each subject at the “missing beat” positions (i.e., the temporal location of strong beat where no physical stimuli were present; Fig. 5, red X'). These time points occurred either 0.25 or 0.5 s after a previous auditory stimulus. To control for carryover effects of evoked responses to previous stimuli, we used the RAND condition (in which there was no expected pulse) as follows: we selected time points that were (1) in the RAND condition where no physical stimulus was present and (2) were at similar distances from previous sounds as the missing beat positions in the MP1 condition (i.e., 0.25 and 0.5 s from the onset of the auditory stimulus). This equated any possible effects of responses to previous sounds on the response at time “zero” between the syncopated and RAND conditions, and enabled us to rule out the possibility that phase locking observed in the syncopated conditions was merely a reflection of carryover responses to previous sounds. For all three conditions, we tested the mean-phase consistency across subjects using the Rayleigh test at each MEG sensor, and corrected for multiple comparisons of the number of MEG sensors using false discovery rate (FDR).

Behavioral tapping paradigm. Following the MEG scans, participants performed a behavioral tapping task to assess their subjective perception of a pulse in syncopated rhythms. This task was performed outside the MEG immediately after the main experiment, and not as part of the main experiment, so as not to contaminate the MEG data with motor activity from the tapping. We used 10 different pulseless syncopated patterns, previously used by Chapin et al. (2010) and Velasco and Large (2011). Each pattern was repeated eight times consecutively (total, 32 s) and

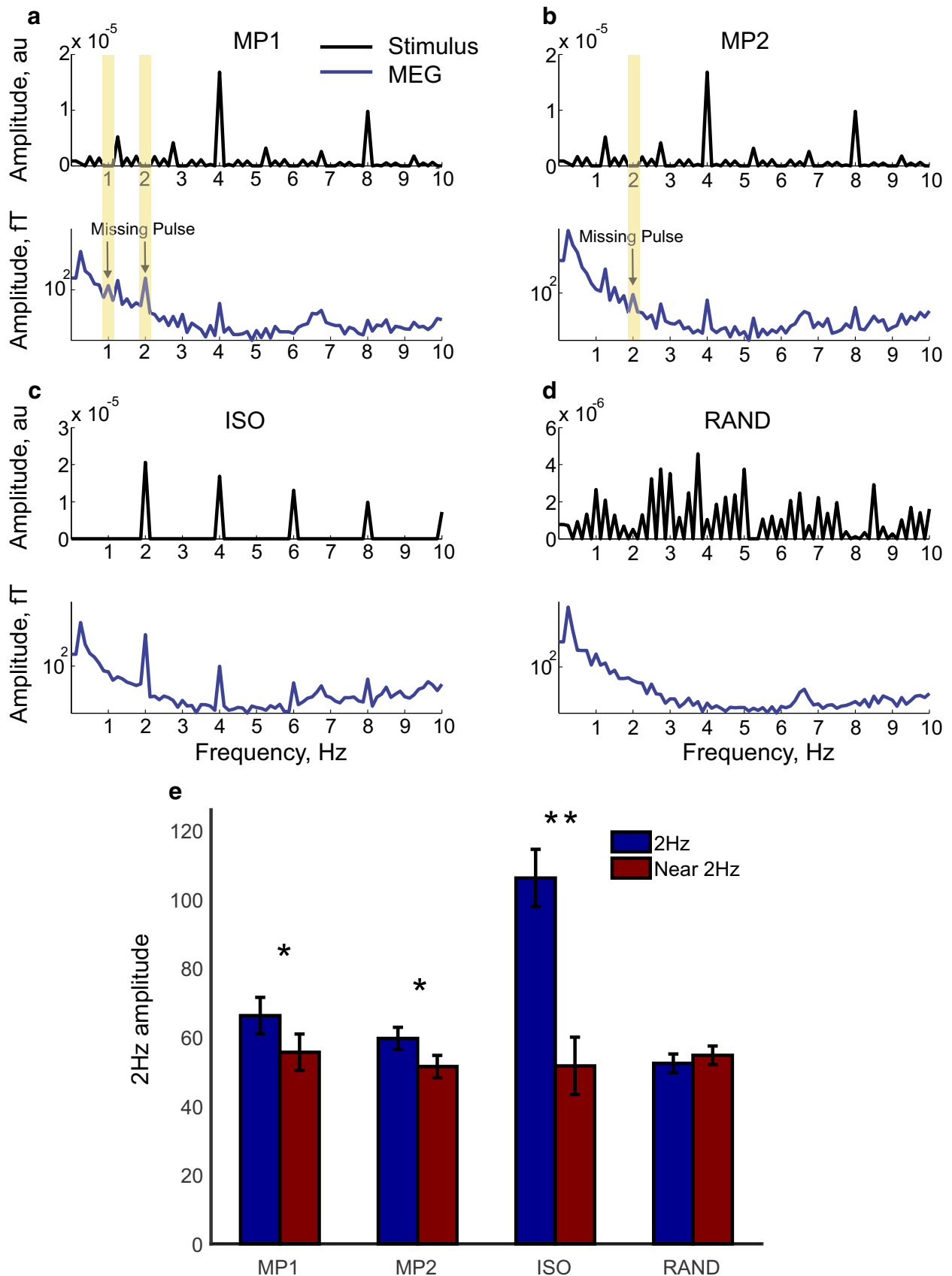


Figure 2. Stimulus and MEG spectrum. *a–d*, The modulation spectra of each stimulus (top) and of the neural response it generated, averaged over 10% of the sensors with the highest 2 Hz amplitude (bottom). *a, b*, Syncopated patterns MP1 and MP2. There is no visible 2 Hz peak in the stimulus spectrum. However, pulse-frequency peaks (black arrows) are observed in the MEG spectrum for both syncopated rhythms. *c*, ISO condition: 2 Hz peaks and its harmonics are clear in both stimulus and MEG spectrum. *d*, RAND pattern: 2 Hz peaks are missing in both the stimulus and the MEG spectrum. *e*, Amplitude at 2 Hz versus the frequencies around it for each condition. Neural data are averaged over all MEG sensors. In the ISO and both (Figure legend continues.)

participants were instructed to listen to the auditory patterns and start tapping on a touch pad in a periodic manner as soon as they perceive a constant pulse from the complex rhythms. To evaluate the robustness of their pulse perception, we evaluated, using circular statistics, how long it took each participant to start tapping [time to tap (TTT)] as well as the temporal precision of their taps. For each subject in each trial, the event times closest to the tapping times were used to convert all taps of all

subjects into a sequence of phases $\phi = 2\pi \frac{t_{event} - t_{tap}}{IOI}$, where the IOI of the metronome time was set to 60/tempo (Large and Gray, 2015). The circular mean for each trial was then computed. The length of the mean vector, r , is the synchronization coefficient, which is a measure of tapping variability (1: no variability; 0: maximum variability). The angle φ of the mean vector shows the relation to the pulse of the rhythmic pattern. $\varphi > 0$ indicates that taps came earlier than the pulse; $\varphi < 0$ indicates that taps came later than the pulse (for additional details of the calculation of circular mean and mean vector, see Large et al., 2015).

Tapping data were not available from one of the subjects due to technical problems. Thus, this analysis was based on the remaining 10 subjects.

Experiment 2

Due to technical problems, we were unable to perform reliable source localization for the data collected in Experiment 1. NRT hypothesized that the two interacting systems involved in generating the missing-pulse responses correspond to sensory and motor networks (Large et al., 2015). To empirically test the predictions of the model as to which brain regions produced the pulse-frequency peaks observed in Experiment 1, we repeated the same paradigm with a different group of subjects, in another MEG scanner. Our analysis focused specifically on contrasting the neural responses in the auditory cortex versus motor-related regions (motor cortex and SMA).

Participants. Eight subjects (three males; mean age, 26 years; SD, 4.17 years) were paid to participate in the study and signed a consent form approved by the Bar Ilan University internal ethics committee. Participants had ≤ 1 year of musical training (in their childhood), had normal hearing, and did not suffer from any neurological disorders.

MEG recordings. Magnetic brain responses were measured with a whole-head helmet-shaped biomagnetometer (4D-Neuroimaging). The sensor array consisted of 248 superconducting magnetometers. A head-position indicator using five coils attached to the scalp provided exact information on the position of the head relative to the sensor array before and after the measurement. Coil positions were determined in relation to external anatomical landmarks. The head shape and coil position were digitized using a Pollhemus Fasttrak digitizer. Brain signals were recorded at a sampling rate of 1017.25 Hz and bandpass filtered at 0.1–400 Hz.

Stimuli and experimental procedure. We used the same stimuli and paradigm as in Experiment 1. However, to assist analysis of the MEG data in source space, auditory and motor localizer tasks were added after the main experiment.

Localizer tasks. The auditory localizer task consisted of passive listening to drumbeat tones presented at random times with an interval of 500 ± 200 ms. The duration of the auditory localizer task was ~ 1.5 min. For the motor localizer task, subjects were instructed to tap repeatedly at a comfortable rate using their index finger for ~ 1.5 min for each hand. Tapping was performed on flat-surfaced custom-made optic device, which did not produce any auditory sounds, and allowed recording the precise tap times for use in subsequent segmentation of the MEG data. The localizer was performed separately for the right and left hands.

Source localization. For source localization of the functional MEG data, we coregistered the head-shape digitization from the MEG with the anatomical MRI of the participant. When an anatomical MRI was not available (three participants) a template MRI (MNI305) was used and

scaled for coregistration with the digitized head shape of the subject. Coregistration used three anatomical landmarks on the subject's head: the left preauricular point (PA), the right PA, and the nasion. Then, we created a Nolte model (Nolte, 2003) hull according to the brain surface of the MRI data using Brainhull software (National Institute of Mental Health, Bethesda, Maryland) and used this hull as a model for source localization.

Brain activity was localized using a synthetic aperture magnetometry (SAM) beamformer (Robinson and Vrba, 1999), which derived a single optimal dipole orientation vector and amplitude for each target voxel that produced the maximal power-to-noise output ratio. We only used the amplitude for further analysis. Voxels were placed on a $5 \times 5 \times 5$ mm rectangular grid. Only voxels within the hull were considered.

Regions of interest selection. Statistical analysis of the power at the pulse frequency was performed in six independently selected regions of interest (ROIs) focusing on bilateral auditory and motor cortices and the SMA. ROIs for the auditory and motor cortices were identified using data from two independent localizer procedures. To this end we calculated the average response in each task [event-related fields (ERFs)] locked either to the onset of the sounds (auditory localizer) or the time of taps (motor localizer). We selected the time window of maximal ERF for each participant to use for ROI localization. Beamformer weights for each voxel were derived from the covariance matrix of the raw localizer data filtered at 1–40 Hz, and these weights were multiplied by the mean ERF in the selected time window to estimate the source currents at each voxel. After aligning the subjects' anatomical and functional data to Talairach coordinates, we chose the 20 voxels exhibiting the strongest responses in each of the localizer tasks within each hemisphere. For the SMA ROI, we used the Human Motor Area Template, a composite atlas based on the meta-analysis of 126 motor-based fMRI studies (Mayka et al., 2006; <http://lrnlab.org>). In total, six ROIs were selected for each subject in the left and right auditory and motor cortices and the SMA, which were used to constrain the statistical analysis of the data in the main experiment.

Localization of pulse frequency. MEG signals were segmented into 32 s epochs corresponding to the passive listening periods and divided into four conditions (ISO, MP1, MP2, RAND). Beamformer weights were calculated for each voxel from the covariance matrix of the raw single-trial data filtered at 0.1–35 Hz to avoid slow drifts or high-frequency noise. Next, the average time course for each condition was multiplied by these weights to obtain the time course of neural responses at each voxel and calculate the spectrum of this response, as in Experiment 1. We normalized the entire spectrum for each voxel by its mean power at 80–120 Hz to avoid bias due to different noise levels at each source location (Shapira Lots et al., 2016). The whole-brain analysis was used to obtain a qualitative description of the regions where the strongest missing-pulse response was observed. However, statistical analysis focused on the six independently selected ROIs, in the bilateral auditory and motor cortices and the SMA. The existence or lack of a missing-pulse response in each of these ROIs was determined using the nonparametric Wilcoxon signed-rank test between the amplitude of the 2 Hz missing-pulse frequency versus the average of amplitude at surrounding frequencies (1.875 and 2.125 Hz). This analysis was conducted on the average spectra across both syncopated stimuli (MP1 and MP2).

Results

Tempo task

The tempo task was designed to keep the subjects attentive to the auditory stimuli and to its timing without explicit instruction to extract the pulse from the syncopated rhythms. All subjects, in both experiments, performed well in the task (mean, 83% correct; SD, 8%) indicating that they indeed attended to the auditory stimuli. There was no significant correlation between the results in the tempo task and the TTT or the synchronization coefficient of the subjects in the subsequent tapping task, although the two subjects who showed the worst performance in the tempo task also performed poorly in the tapping task.

←

(Figure legend continued.) syncopated conditions there was a significant difference between 2 Hz power and the frequencies around it, indicating a peak at the pulse frequency (Wilcoxon signed-rank test, $*p < 0.05$, $**p < 0.01$). In the RAND condition, there was no significant peak at 2 Hz. Error bars indicate SEM.

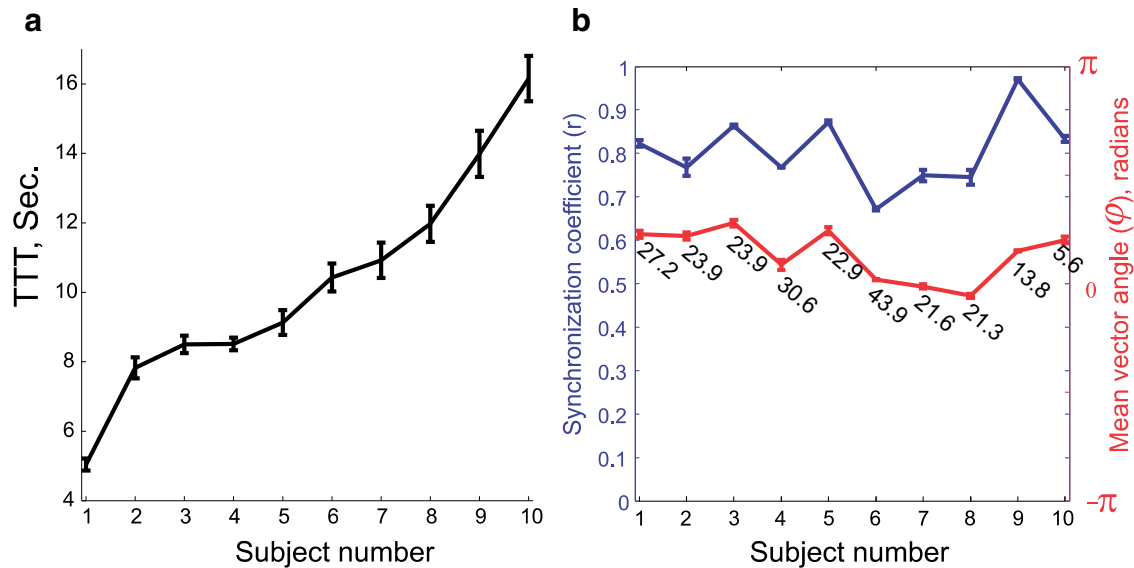


Figure 3. Behavioral performance in the tapping task. *a*, TTT, the time elapsed from the beginning of the stimulus until subjects started tapping at the perceived pulse frequency. *b*, Blue, Synchronization coefficient for each subject. All subjects eventually managed to tap in synchrony with the beat. Red, Angle of the mean vector. Almost all subjects showed positive values indicating that tapping came earlier than the strong-beat positions. Error bars indicate SEM. The numbers on the bottom of the red trace indicate the average number of taps for each subject over the 10 syncopated rhythms during the behavioral task.

MEG results—Experiment 1

The main goal of this study was to test whether we could identify the emergence of the pulse in the dynamics of neural activity despite its absence in the acoustics. Figure 2*a* shows a comparison of the frequency content of each of the four stimuli used (top) and the spectrum of the MEG responses to each stimulus (bottom), averaged over all subjects.

The spectrum of the isochronous stimulus contains power at 2 Hz and its harmonics, and this pattern is faithfully represented in the neural response. The spectra of the two syncopated stimuli (MP1, MP2) have peaks at many frequencies (e.g., 0.75, 1.25, 2.75, and 4 Hz), but contain no energy at 2 Hz, which is the expected pulse frequency. The MEG-recorded neural responses to these stimuli show peaks at all the frequencies contained in the stimuli. However, in addition, clear 2 Hz peaks are observed for both MP1 and MP2, despite the absence of this frequency in the stimulus spectrum. Moreover, in MP1, a peak was seen at 1 Hz as well (a subharmonic that can also be perceived as the pulse in some cases; Parncutt, 1994). These results demonstrate the brain’s capacity for generating an internal representation of the perceived pulse, above and beyond the acoustic of the stimulus. In the RAND condition, the stimulus spectrum contains several small peaks corresponding to peaks in the stimulus spectrum, but does not show any peak at 2 Hz, nor at any other frequency not present in the stimulus itself.

Correlations with tapping to the beat

Pulse perception for complex rhythms is not necessarily ubiquitous, but may vary across individuals. Therefore, we conducted an independent assessment of our participants’ ability to extract the pulse from syncopated stimuli in a tapping task, performed outside the MEG immediately after the main experiment. There was substantial variability in the time it took each participant to perceive the pulse and start tapping (TTT; Fig. 3*a*), with some tapping within the second repetition and others requiring three or four repetitions of the rhythm (>10 s) to start tapping. However, once they started tapping, most subjects successfully syn-

chronized to the stimulus and tapped at either 1 or 2 Hz, both of which could be induced as the pulse of these stimuli (Honing, 2013; Large et al., 2015). Individual participants did not necessarily exhibit a clear “preference” for tapping at one of these frequencies for all 10 stimuli, and in some instances started tapping at one frequency (e.g., 2 Hz) and shifted to another frequency (e.g., 1 Hz) midtrial. This variability in pulse perception is consistent with existing literature demonstrating that the same stimulus can induce the perception of a pulse at different (integer-related) frequencies across and even within individuals (for review, see Repp and Su, 2013). Variability in tap rate was accounted for when calculating the synchronization coefficients and the angle of the mean vector, which ranged between $0.67 < r < 0.97$ (Fig. 3*b*, blue trace) and between $-0.17 < \varphi < 0.87$ (Fig. 3*b*, red trace), respectively. There was no significant correlation between the TTT and the synchronization coefficients of the subjects, indicating that once participants perceived the pulse and started tapping, they did so accurately.

The variability in the time it takes individuals to perceive a pulse provides an opportunity to test whether the emergence of a neural response at the pulse rate was uniform across all participants or whether it covaried with their subjective perception of a pulse. Figure 4 shows a comparison of the MEG spectrum in the syncopated conditions for two example fast and slow pulse perceivers, as assessed by their TTT values. Indeed, the fast pulse perceiver showed the emergence of power at the 2 Hz pulse frequency, whereas this peak is absent in the spectra of the slow pulse perceiver, supporting the notion that the 2 Hz peak is related to the perception of a pulse rather than the acoustics of the stimulus. To quantify this effect, we calculated Spearman’s rank correlation between the TTT and the 2 Hz amplitude in the neural signal, averaged over all MEG sensors. We found a strong negative correlation between the TTT and the 2 Hz power ($r_s = -0.82$, $p < 0.01$; Fig. 4*c*), such that subjects who started to tap to the pulse sooner had higher amplitude at 2 Hz when passively listening to syncopated stimuli, whereas 2 Hz amplitude was lower in participants who took longer to detect a pulse. TTT was not signifi-

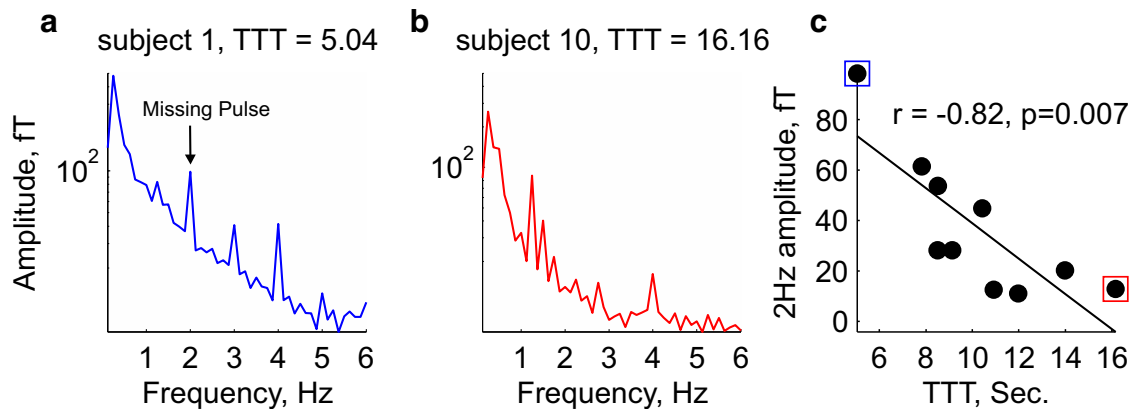


Figure 4. MEG spectrum for individual subjects. The spectrum is shown averaged over both syncopated conditions and over seven channels (5% of the total number of channels) showing the highest 2 Hz power. **a**, A subject with low average TTT value, who perceived the pulse of the syncopated rhythms early on during the behavioral task. **b**, A subject with high TTT value. It took the subject >15 s to perceive the pulse. The MEG spectrum of the early pulse perceiver shows a clear 2 Hz peak that matches the expected pulse frequency of the stimulus while the late pulse perceiver failed to generate such oscillations. Note that both subjects showed peaks at frequencies present in the stimulus itself (e.g., 1.25 and 4 Hz). **c**, Correlation between 2 Hz amplitude and TTT for the two syncopated conditions. Subjects that showed higher 2 Hz amplitude during the MEG scans in the syncopated conditions perceived the pulse of the syncopated patterns and started tapping earlier during the subsequent behavioral task. Blue and red squares indicate the subjects shown in **a** and **b** respectively.

cantly correlated with 2 Hz amplitude in the ISO or RAND condition ($p > 0.5$ for both).

Phase locking to the missing pulse

Thus far we have inferred the existence of a pulse response by looking at the spectra of the neural response, averaged across the entire stimulus. However, the special nature of the syncopated stimuli used here allows us to specifically test the hypothesis that pulse induction is manifest in entrainment/phase-locking of low-frequency neural activity to the pulse rhythm (Large et al., 2015). We use the term “entrainment” to describe the synchronization of neural responses to an external (or perceptual) rhythm, which in this case is the “pulse” rhythm. According to this hypothesis, we would expect to find increased low-frequency phase locking at strong-beat positions within syncopated sequences. The syncopated stimuli used here were constructed specifically so that individual sounds coincided with strong-beat positions only half the time, whereas in the other half of the strong-beat positions no sounds occurred. Thus, in this analysis we focused only on the neural responses time-locked to strong-beat positions where no sounds occurred (Fig. 5*a*, top, red X’s), thus avoiding contamination with sensory-evoked potentials. The bottom panels in Figure 5*a* show the average sounds wave-locked to these positions, verifying that these were indeed locked to “quiet” periods within the sequence, and shared similar acoustic properties with control time periods selected from the RAND condition.

Figure 5*b* shows the grand average for neural responses time-locked to these positions for the two syncopated conditions and for the RAND condition. For both syncopated conditions, a negative deflection is observed locked to the soundless strong-beat positions ($t = 0$), whereas in the RAND condition the responses are flat.

To quantify this response, we filtered the single-trial responses between 0.5 and 3 Hz and extracted the momentary phase at this position separately for each trial, sensor, and participant and calculated the mean phase across trials. We then tested the phase consistency across participants using the Rayleigh test. A distribution of the mean phase across subjects is shown in Figure 5*c* for one representative MEG sensor. Approximately 30% of the MEG sensors showed significant phase locking at the time of the “soundless strong-beat position” in both syncopated conditions

(50 sensors for MP1; 51 for MP2; $p < 0.05$, FDR corrected). The topographical distribution of sensors with significant phase locking corresponded to the typical auditory topography (Fig. 5*d*). No significant phase locking was found in the RAND condition, indicating that the effects found for the syncopated stimuli do not result merely from a carryover of evoked responses to preceding sounds (see Materials and Methods).

Where is the neural correlate of the missing pulse generated?—Experiment 2

After establishing the emergence of a pulse-related neural dynamic dissociated from simple evoked responses, we ask which brain regions are involved in generating this missing-pulse response profile. To test the origin of the missing-pulse response, we repeated the experiment with a new set of participants, in another MEG facility. The whole-brain analysis revealed that the strongest missing-pulse response was found in two clusters localized to the right and left STGs, with the response stronger on the right (Fig. 6*c*). These clusters overlapped with the clusters detected in the auditory localizer task. Statistical analysis was performed in six independently selected ROIs corresponding to the bilateral auditory and motor cortices and the SMA. The spectra of the neural response to syncopated stimuli in each ROI are shown in Figure 6*a*, revealing a missing-pulse response at 2 Hz only in the STG ROIs. Nonparametric statistical analysis confirmed that the 2 Hz missing-pulse response was significantly higher than neighboring nonpulse frequencies (averaged over 1.875 and 2.125 Hz) in the right STG (Wilcoxon signed-rank test, $p < 0.05$) and the left STG showed a trend toward significance ($p = 0.071$), whereas the 2 Hz response was not significant in the motor (right: $p > 0.36$; left: $p > 0.18$) or SMA ROIs (right: $p > 0.52$; left: $p > 0.26$). Thus, we replicated the missing-pulse phenomenon found in Experiment 1, and found that a correlate of this effect was primarily localized to the right STG.

Note that the spectra in all ROIs, including motor regions, showed peaks at frequencies that exist in the stimulus itself (which was the case for the isochronous stimulus as well; data not shown), consistent with previous studies suggesting activation in motor-related regions in response to auditory stimuli (Grahn and Bret, 2007; Grahn and Rowe, 2009; Patel and Iversen, 2014; Arnal et al., 2015).

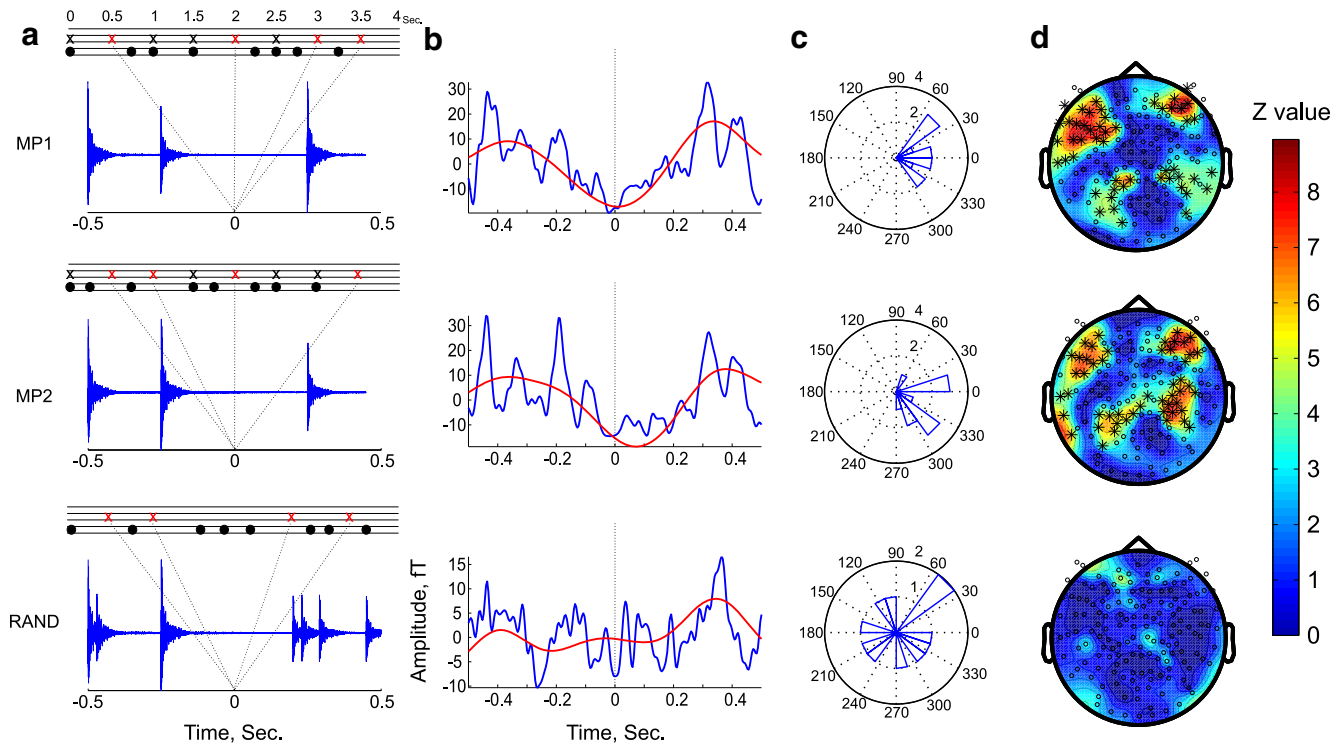


Figure 5. Phase locking to the missing pulse for the two syncopated rhythms (MP1 and MP2; top two rows) and the RAND condition (bottom row). **a**, Illustration of the soundless strong-beat positions. Black dots mark the locations of each physical sound and, for the two syncopated rhythms, X's mark the strong-beat positions. Red X's denote strong-beat positions where no sounds were present, which served as $t = 0$ in this analysis. The bottom panels show averaged sound files time-locked to these positions, verifying that these were indeed quiet times. Similar soundless periods were selected from the RAND condition (depicted by the red X's), matched on their distance from previous sounds so as to control for carryover effects of evoked responses from previous sounds. **b**, MEG signals averaged over all participants for one representative MEG sensor. The signals were time-locked to the position of soundless strong beats ($t = 0$) in the syncopated conditions, or control epochs in the RAND condition, filtered between 1 and 40 Hz (blue) and between 0.5 and 3 Hz (red). **c**, Phase histograms at the time of the expected pulse, across all participants for the MEG sensor shown in **b**. **d**, Topography of the phase-locking values across MEG sensors. The colors indicate the Z value of the Rayleigh test. Significant sensors are indicated by an asterisk ($p < 0.05$, FDR corrected).

Discussion

The current results are in line with theoretical predictions of NRT (Large and Snyder, 2009) associating pulse perception with emergence of neural phase locking at the pulse frequency. NRT predicts that, due to higher-order resonances, when neural oscillators are driven by an acoustic stimulus, oscillations may emerge at favored frequency ratios (harmonics, subharmonics, and integer ratios), not always present in a linear frequency analysis of the acoustics (Velasco and Large, 2011). Empirical findings consistent with this model show enhancement of pulse-related frequencies in EEG-recorded steady-state evoked potentials (SS-EPs) compared with nonpulse frequencies (Nozaradan et al., 2012) and its correlation with behavioral measures of pulse perception (Nozaradan et al., 2016). However, in those studies, most stimuli contained the pulse frequency in their modulations spectra, allowing alternative interpretations suggesting that SS-EP modulations reflect enhancement of evoked responses to on-beat sounds (Potter et al., 2009; Schaefer et al., 2011). That said, at least one stimulus used in those studies could be considered a missing-pulse rhythm, supporting the current findings.

Our study comprehensively tests the predictions of NRT using missing-pulse rhythms to rule out the following two alternative explanations: (1) that apparent entrainment of SS-EPs merely reflects overlapping transient responses to individual sounds (Tremblay et al., 2004; Sussman et al., 2008) or (2) that apparent entrainment of SS-EPs arises from linear neural frequency analysis, such as a bank of bandpass filters (Todd, 1999; Tomic and Janata, 2008). In both cases the spectra of the neural response to missing-pulse rhythms should

mirror the modulations spectrum of the stimuli, and would not contain energy at the pulse frequency. However, our results show clear missing-pulse responses—both in the spectrum of the neural response and in phase locking at the time of unsounded beats—findings that cannot be explained by simple linear neural dynamics.

That said, current findings do not unequivocally prove that the observed pulse response results from synchronization of endogenous neural oscillation. An alternative model also consistent with the current results is that of neural time delay, which predicts responses at subharmonics of the dominant 4 Hz frequency in the stimuli (Scheirer, 1998; Eck, 2006). Additional experiments are required to determine which model is superior in describing the underlying computations involved in neural encoding of complex rhythms.

Interestingly, the phenomenon identified here—perceptual and neural responses to a missing pulse—is analogous to the “case of the missing fundamental,” where the perceived pitch of harmonic tones correspond to the fundamental frequency (F_0) even if that frequency is physically removed from the stimuli (Licklider, 1954). Although the mechanistic infrastructure underlying the two phenomena may differ (Meddis and O'Mard, 2006; Kerr et al., 2013; Huang and Rinzel, 2016), in both cases the most salient percept corresponds to a frequency missing in an objective frequency analysis of stimulus acoustics, demonstrating a disparity between sensation and perception in audition.

Phase locking to the missing pulse

The analysis of low-frequency phase locking at soundless strong-beat positions enabled us to further disentangle the manifestation

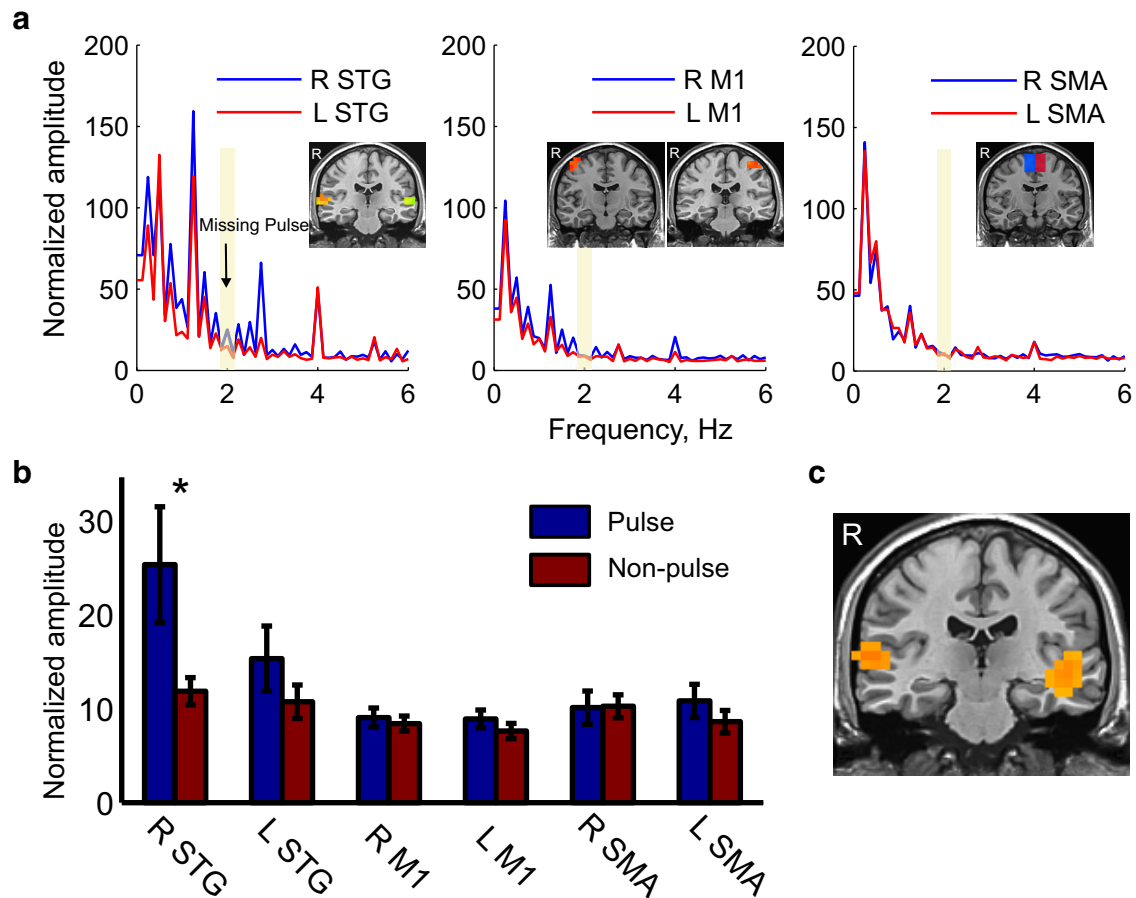


Figure 6. Localization of the missing-pulse response. *a*, Spectra of the neural response in response to the two syncopated rhythms in the right (blue) and left (red) hemisphere for each ROI. Inserts indicate the anatomical locations of the voxels used for each ROI (see Materials and Methods). *b*, Statistical comparison of 2 Hz amplitude (blue) versus amplitude in adjacent frequencies (red) in each ROI. Error bars indicate SEM. *c*, Whole-brain analysis of the missing-pulse response. The two clusters with the strongest missing-pulse responses were localized in the right and left STGs.

of neural synchronization from stimulus-driven evoked responses. It has been proposed that slow cortical oscillations can be used to tune neural excitability to the timing of upcoming events and thus form the basis for predictive timing (Schroeder and Lakatos, 2009; Arnal and Giraud, 2012). Low-frequency phases can be modulated by top-down processes to enhance stimulus processing (Buschman and Miller, 2007; Lakatos et al., 2008, 2013, 2016; Arnal et al., 2015) and has been linked to increased perceptual sensitivity (Large and Jones, 1999; Ng et al., 2012; Zoefel and VanRullen, 2015). The occurrence of phase locking in the absence of physical stimuli suggests the influence of intrinsic processes related to extracting the underlying temporal structure of the sequence on local neuronal excitability.

Variability in pulse perception

Pulse perception is not ubiquitous within the population (Sowiński and Dalla Bella, 2013), with an extreme case being beat-deaf individuals who fail in perceiving the pulse of syncopated rhythms (Phillips-Silver et al., 2011). Here, we observe variability in the time it takes for pulse induction, reflected in participants' TTT, which was correlated with the magnitude of the neural missing-pulse response. While TTT remains an indirect measure of pulse induction, it probably gives us an upper-bound estimate, especially since once participants started tapping, they did so accurately, indicating that no one started tapping before perceiving the pulse. Thus, our results suggest a direct link between the neural pulse-frequency response and subjective pulse perception.

These results align well with findings that activity in pulse-related regions does not occur “automatically” (Chapin et al., 2010; Doelling and Poeppel, 2015). The current findings also predict that beat-deaf individuals will not exhibit a missing-pulse response in their neural activity. This correlation may be interpreted either as a stronger pulse response, or as an earlier onset of the pulse response in people faster at detecting the pulse. Unfortunately, attempts to test these two options by studying the development of pulse responses over time did not yield substantial results.

Another source of potential variability across (and within) individuals is the specific frequency induced as the pulse. For the stimuli used here, a 2 Hz pulse is typically perceived. However, a 1 Hz pulse can also be induced (Large et al., 2015). Indeed, in the behavioral session participants tapped at both rates for different stimuli, without any discernable individual preference. Moreover, the neural response to stimulus MP1 exhibited a peak at 1 Hz, in addition to the 2 Hz peak. It is well established that pulse perception is dynamic and flexible, and the perceived pulse frequency for a given stimulus can vary across and even within individuals (for review, see Repp and Su, 2013), and can be intentionally changed (Large et al., 2002; Nozaradan et al., 2011). Thus, another prediction to be tested in future studies is whether the dominant missing-pulse frequency in the neural signal reflects the specific pulse frequency induced in a given trial. Unfortunately, given the difference in stimuli used in the behavioral and MEG sessions here, and the fact that these were tested sepa-

rately, the current data are not suitable for addressing this question.

What is the source?

The missing-pulse response was predominantly localized to the right auditory cortex. This is consistent with the asymmetrical sampling in time framework, which suggests that the auditory cortices in the two hemispheres differ in their sensitivities to temporal information (Poehppel, 2003; Giraud et al., 2007). The right hemisphere is proposed to be more sensitive to slow time scales, relevant for musical pulse, whereas the left hemisphere shows advantages for processing rapid sound variations.

Several studies suggest that the motor system is also involved in pulse perception (Merchant et al., 2015). For example, activation of motor areas bilaterally has been reported during rhythm perception in the absence of movement (Fujioka et al., 2012; Grahn and Bret, 2007; Michaelis et al., 2014; Patel and Iversen, 2014). However, in the current study, we did not detect significant pulse-related responses in motor areas. While this does not necessarily contradict existing literature showing that the motor system plays a role in pulse perception, it does leave open the question of the specific nature of the neural dynamics and representation of the pulse within these regions.

The NRT framework, which describes two interacting oscillatory networks intended to model auditory and motor dynamics, may also need to be modified and updated in future iterations of the model, based on the current findings. Specifically, the dynamic interactions predicted by the model may occur primarily in subnetworks within a single area (e.g., the auditory cortex), and not between spatially distinct brain regions as originally proposed. Such relabeling of the coupled networks would not change the core predictions of the model as to the dynamic pattern of neural responses, but would alter predictions as to where these responses would be observed. Additional research is ongoing to clarify this point before updating the NRT model. However, as we see it, one of the strengths of this study is the empirical testing of specific predictions put forth by a theoretical model, which may well result in updating some aspects of the model to improve its biological plausibility.

Conclusions

We demonstrate the capacity of the auditory cortex to phase-lock to the pulse of complex rhythms, and successfully dissociate this activity from evoked responses. Moreover, we show that this phenomenon is not seen in individuals who have difficulty perceiving the pulse, linking it directly to pulse perception. These findings bear significance not only for music perception, but for speech processing as well, which also contains temporal regularities at similar time scales (Zion Golumbic et al., 2012). Entrainment of cortical activity to rhythmic stimuli has been suggested as a mechanism for enhancing processing efficiency; decoding high-level information and directing selective attention in a noisy environment (Schroeder and Lakatos, 2009; Zion Golumbic et al., 2013; Kayser et al., 2015; Lakatos et al., 2016; Zoefel and VanRullen, 2016). Use of complex yet precisely controlled stimuli, coupled with empirical testing of specific theoretical models, provides a unique opportunity to study such mechanisms, leading to a better understanding of neural dynamics underlying the processing of continuous naturalistic sounds.

Notes

Supplemental material for this article, including the audio files of stimuli used in the experiment, is available at <http://www.golumbiclab.org/mpstimuli>. This material has not been peer reviewed.

References

- Arnal LH, Giraud AL (2012) Cortical oscillations and sensory predictions. *Trends Cogn Sci* 16:390–398. [CrossRef Medline](#)
- Arnal LH, Doelling KB, Poeppel D (2015) Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cereb Cortex* 25:3077–3085. [CrossRef Medline](#)
- 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* 45:62–71. [CrossRef Medline](#)
- Bouwer FL, Honing H (2015) Temporal attending and prediction influence the perception of metrical rhythm: evidence from reaction times and ERPs. *Front Psychol* 6:1094. [CrossRef Medline](#)
- Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862. [CrossRef Medline](#)
- Chapin HL, Zanto T, Jantzen KJ, Kelso SJ, Steinberg F, Large EW (2010) Neural responses to complex auditory rhythms: the role of attending. *Front Psychol* 1:224. [CrossRef Medline](#)
- Chen JL, Penhune VB, Zatorre RJ (2009) The role of auditory and premotor cortex in sensorimotor transformations. *Ann N Y Acad Sci* 1169:15–34. [CrossRef Medline](#)
- de Cheveigné A, Simon JZ (2007) Denoising based on time-shift PCA. *J Neurosci Methods* 165:297–305. [CrossRef Medline](#)
- Doelling KB, Poeppel D (2015) Cortical entrainment to music and its modulation by expertise. *Proc Natl Acad Sci U S A* 112:E6233–E6242. [CrossRef Medline](#)
- Eck D (2006) Identifying metrical and temporal structure with an autocorrelation phase matrix. *Music Perception* 24:167–176. [CrossRef](#)
- Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *J Neurosci* 32:1791–1802. [CrossRef](#)
- Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RS, Laufs H (2007) Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56:1127–1134. [CrossRef Medline](#)
- Grahn JA, Brett M (2007) Rhythm and beat perception in motor areas of the brain. *J Cogn Neurosci* 19:893–906. [CrossRef Medline](#)
- Grahn JA, Rowe JB (2009) Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J Neurosci* 29:7540–7548. [CrossRef Medline](#)
- Honing H (2013) Structure and interpretation of rhythm in music. *Psychol Music* 3:369–404. [CrossRef](#)
- Huang C, Rinzel J (2016) A neuronal network model for pitch selectivity and representation. *Front Comput Neurosci* 10:57. [CrossRef Medline](#)
- Kayser SJ, Ince RA, Gross J, Kayser C (2015) Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal α . *J Neurosci* 35:14691–14701. [CrossRef Medline](#)
- Kerr RR, Burkitt AN, Thomas DA, Gilson M, Grayden DB (2013) Delay selection by spike-timing-dependent plasticity in recurrent networks of spiking neurons receiving oscillatory inputs. *PLoS Comput Biol* 9:e1002897. [CrossRef Medline](#)
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–113. [CrossRef Medline](#)
- Lakatos P, Musacchia G, O’Connell MN, Falchier AY, Javitt DC, Schroeder CE (2013) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77:750–761. [CrossRef Medline](#)
- Lakatos P, Barczak A, Némotín SA, McGinnis T, Ross D, Javitt DC, O’Connell MN (2016) Global dynamics of selective attention and its lapses in primary auditory cortex. *Nat Neurosci* 19:1707–1717. [CrossRef Medline](#)
- Langner G (1992) Periodicity coding in the auditory system. *Hear Res* 60:115–142. [CrossRef Medline](#)
- Large EW (2008) Resonating to musical rhythm: theory and experiment. In: *Psychology of time* (Grondin S, ed), pp 189–232. Bingley, UK: Emerald Group.
- Large EW, Gray PM (2015) Spontaneous tempo and rhythmic entrainment in a bonobo (*Pan paniscus*). *J Comp Psychol* 129:317–328. [CrossRef Medline](#)
- Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. *Psychol Rev* 106:119. [CrossRef](#)

- Large EW, Snyder JS (2009) Pulse and meter as neural resonance. *Ann N Y Acad Sci* 1169:46–57. [CrossRef Medline](#)
- Large EW, Fink P, Kelso JA (2002) Tracking simple and complex sequences. *Psychol Res* 66:3–17. [CrossRef Medline](#)
- Large EW, Herrera JA, Velasco MJ (2015) Neural networks for beat perception in musical rhythm. *Front Syst Neurosci* 9:159. [CrossRef Medline](#)
- Licklider JCR (1954) “Periodicity” pitch and “place” pitch. *J Acoust Soc Am* 26:945. [CrossRef](#)
- Mayka MA, Corcos DM, Leurgans SE, Vaillancourt DE (2006) Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *Neuroimage* 31:1453–1474. [CrossRef Medline](#)
- Meddis R, O’Mard LP (2006) Virtual pitch in a computational physiological model. *J Acoust Soc Am* 120:3861–3869. [CrossRef Medline](#)
- Merchant H, Grahm J, Trainor L, Rohrmeier M, Fitch WT (2015) Finding the beat: a neural perspective across humans and non-human primates. *Philos Trans R Soc Lond B Biol Sci* 370:20140093. [CrossRef Medline](#)
- Michaelis K, Wiener M, Thompson JC (2014) Passive listening to preferred motor tempo modulates corticospinal excitability. *Front Hum Neurosci* 8:252. [CrossRef Medline](#)
- Moelants D (2002) Preferred tempo reconsidered. In: *Proceedings of the 7th International Conference on Music Perception and Cognition*. Sydney, Australia, July.
- Nettl B (2000) An ethnomusicologist contemplates universals in musical sound and musical culture. In: *The origins of music* (Wallin N, Merker J, Brown S, eds), pp 463–472. Cambridge, MA: MIT.
- Ng BS, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* 32:12268–12276. [CrossRef Medline](#)
- Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol* 48:3637–3652. [CrossRef Medline](#)
- Nozaradan S, Peretz I, Missal M, Mouraux A (2011) Tagging the neuronal entrainment to beat and meter. *J Neurosci* 31:10234–10240. [CrossRef Medline](#)
- Nozaradan S, Peretz I, Mouraux A (2012) Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J Neurosci* 32:17572–17581. [CrossRef Medline](#)
- Nozaradan S, Peretz I, Keller PE (2016) Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. *Sci Rep* 6:20612. [CrossRef Medline](#)
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869. [CrossRef Medline](#)
- Parncutt R (1994) A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept* 11:409–464. [CrossRef](#)
- Patel AD, Iversen JR (2014) The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front Syst Neurosci* 8:57. [CrossRef Medline](#)
- Patel AD, Iversen JR, Chen Y, Repp BH (2005) The influence of metricality and modality on synchronization with a beat. *Exp Brain Res* 163:226–238. [CrossRef Medline](#)
- Phillips-Silver J, Toiviainen P, Gosselin N, Piché O, Nozaradan S, Palmer C, Peretz I (2011) Born to dance but beat deaf: a new form of congenital amusia. *Neuropsychologia* 49:961–969. [CrossRef Medline](#)
- Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun* 41:245–255. [CrossRef](#)
- Potter DD, Fenwick M, Abecasis D, Brochard R (2009) Perceiving rhythm where none exists: event-related potential (ERP) correlates of subjective accenting. *Cortex* 45:103–109. [CrossRef Medline](#)
- Repp BH, Su YH (2013) Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon Bull Rev* 20:403–452. [CrossRef Medline](#)
- Robinson SE, Vrba J (1999) Functional neuroimaging by synthetic aperture magnetometry (SAM). In: *Recent advances in biomagnetism*. (Yoshimoto T, Kotani M, Kuriki S, Karibe H, Nakasoto N, eds), pp 302–5. Sendai, Japan: Tohoku UP.
- Schaefer RS, Vlek RJ, Desain P (2011) Decomposing rhythm processing: electroencephalography of perceived and self-imposed rhythmic patterns. *Psychol Res* 75:95–106. [CrossRef Medline](#)
- Scheirer ED (1998) Tempo and beat analysis of acoustic musical signals. *J Acoust Soc Am* 103:588–601. [CrossRef Medline](#)
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18. [CrossRef Medline](#)
- Shapira Lots I, Zeev-Wolf M, Harpaz Y, Abeles M (2016) Source localization scale correction for Beamformer analysis. *J Neurosci Methods* 273:10–19. [CrossRef Medline](#)
- Snyder JS, Large EW (2005) Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Res Cogn Brain Res* 24:117–126. [CrossRef Medline](#)
- Sowiński, Dalla Bella S (2013) Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia* 51:1952–1963. [CrossRef Medline](#)
- Sussman E, Steinschneider M, Gumenyuk V, Grushko J, Lawson K (2008) The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res* 236:61–79. [CrossRef Medline](#)
- Tierney A, Kraus N (2015) Neural entrainment to the rhythmic structure of music. *J Cogn Neurosci* 27:400–408. [CrossRef Medline](#)
- Todd N (1999) Implications of a sensory-motor theory for the representation and segregation of speech. *J Acoust Soc Am* 105:1307. [CrossRef](#)
- Tomic ST, Janata P (2008) Beyond the beat: modeling metric structure in music and performance. *J Acoust Soc Am* 124:4024–4041. [CrossRef Medline](#)
- Tremblay KL, Billings C, Rohila N (2004) Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 15:226–237; quiz 264. [CrossRef Medline](#)
- Velasco MJ, Large EW (2011) Pulse detection in syncopated rhythms using neural oscillators. Paper presented at the 12th Annual Conference of the International Society for Music Information Retrieval. Miami, October.
- Zion Golumbic EM, Poeppel D, Schroeder CE (2012) Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain Lang* 122:151–161. [CrossRef Medline](#)
- Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* 77:980–991. [CrossRef Medline](#)
- Zoefel B, VanRullen R (2015) Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *J Neurosci* 35:1954–1964. [CrossRef Medline](#)
- Zoefel B, VanRullen R (2016) EEG oscillations entrain their phase to high-level features of speech sound. *Neuroimage* 124:16–23. [CrossRef Medline](#)