CrossMark

ORIGINAL ARTICLE

Intracerebral evidence of rhythm transform in the human auditory cortex

Sylvie Nozaradan[1,2,3] · André Mouraux[1] · Jacques Jonas[1,4,5] · Sophie Colnat-Coulbois[7] ·
Bruno Rossion[1,4,6] · Louis Maillard[4,5]

Abstract Musical entrainment is shared by all human cultures and the perception of a periodic beat is a cornerstone of this entrainment behavior. Here, we investigated whether beat perception might have its roots in the earliest stages of auditory cortical processing. Local field potentials were recorded from 8 patients implanted with depth-electrodes in Heschl's gyrus and the planum temporale (55 recording sites in total), usually considered as human primary and secondary auditory cortices. Using a frequency-tagging approach, we show that both low-frequency (<30 Hz) and high-frequency (>30 Hz) neural activities in these structures faithfully track auditory rhythms through frequency-locking to the rhythm envelope. A selective gain in amplitude of the response frequency-locked to the beat frequency was observed for the low-frequency activities but not for the high-frequency activities, and was sharper in the planum temporale, especially for the more challenging syncopated rhythm. Hence, this gain process is not systematic in all activities produced in these areas and depends on the complexity of the rhythmic input. Moreover, this gain was disrupted when the rhythm was presented at fast speed, revealing low-pass response properties which could account for the propensity to perceive a beat only within the musical tempo range. Together, these observations show that, even though part of these neural transforms of rhythms could already take place in subcortical auditory processes, the earliest auditory cortical processes shape the neural representation of rhythmic inputs in favor of the emergence of a periodic beat.

✉ Sylvie Nozaradan
sylvie.nozaradan@uclouvain.be

1 Institute of Neuroscience (Ions), Université catholique de Louvain (UCL), 53, Avenue Mounier, UCL 53.75, 1200 Brussels, Belgium

2 The MARCS Institute, Western Sydney University, Sydney, NSW 2214, Australia

3 International Laboratory for Brain, Music and Sound Research (Brams), Montreal H3C 3J7, Canada

4 Service de Neurologie, Centre Hospitalier Universitaire de Nancy, 54035 Nancy, France

5 CRAN UMR 7039 CNRS Université de Lorraine, 54035 Nancy, France

6 Psychological Sciences Research Institute, Université Catholique de Louvain (UCL), 1348 Louvain-la-Neuve, Belgium

7 Neurosurgery Department, Centre Hospitalier Universitaire de Nancy, 54035 Nancy, France

Introduction

Among all sensory systems, audition excels in encoding the temporal dimension of incoming inputs. The high sensitivity of the auditory system to temporal features of the stimulus has been observed in the earliest electrophysiological recordings (see Erulkar et al. 1968; Fernald and Gerstein 1972), and is equally prominent in behavioral observations of humans and non-human animal species (e.g., Joris et al. 2004 for a review). Importantly, multiple temporal levels characterize acoustic stimuli: while fast pressure variations determine the spectral content of sounds, constituting the fine-structure, the slower modulation of this structure corresponds to the sound envelope.

⚡ Springer

Among the range of frequencies audible for the human ear, sound components above 100 Hz code for a fundamental frequency called *pitch* (see e.g. Smith et al. 2002; Malone and Schreiner 2010 for recent reviews), and the harmonics thereof define fine features, such as timbre or loudness. Between ∼25–125 Hz, envelope modulations convey a perception of roughness. Finally, modulations below ∼10 Hz, which include the average rate of syllables in speech and convey the perception of rhythms and tempo in music (Fraisse 1967; Drake and Botte 1993), appear to play a specific role for auditory communication (Drullman et al. 1994a, b; Shannon et al. 1995) and auditory-motor synchronization (Repp 2005; Repp and Su 2013).

The human brain can faithfully track—i.e., time-lock to—sound envelopes containing slow fluctuations, as revealed by scalp surface and intracranial electroencephalographic (EEG) recordings (Pantev et al. 1988; Picton et al. 1987; Eggermont 2001; Liégeois-Chauvel et al. 2004). The relationship between the sound input and the neural output, defined by a *temporal modulation transfer function* (see Møller 1972 for the first description of the concept in hearing research), can be derived by comparing the two signals, with the aim to uncover the neural mechanisms leading to the emergence of a perceptual representation of the sensory stimulus (Edwards and Chang 2013).

Up to now, electrophysiological investigations of sound envelope tracking have been mostly conducted using either extremely simplified stimuli or complex speech sequences. Studies using simplified stimuli, namely trains of periodic amplitude-modulated acoustic stimuli such as clicks or sinusoidal modulations of a tone, have provided evidence for sound envelope tracking in various human cortical areas, including Heschl's gyrus and the planum temporale, two brain regions typically assumed to correspond, respectively, to the primary and secondary auditory cortices (e.g., Pantev et al. 1988; Picton et al. 1987; Eggermont 2001; Liégeois-Chauvel et al. 2004). Moreover, responses to isochronous sounds have also been observed in the form of modulations of brain oscillations (e.g., beta-band oscillations at ∼20–30 Hz) generated within auditory cortices as reconstructed from scalp magnetoencephalographic recordings, and tracking the incoming sounds over time (Fujioka et al. 2012). Some of these studies have reported time constraints biasing the envelope tracking in these areas towards specific band-pass tuning ranges (e.g. ∼2–5 Hz, see Edwards and Chang 2013, or 4–16 Hz, see Gourévitch et al. 2011). However, whether these findings can be generalized to stimuli with higher temporal complexity remains unclear. Even though speech constitutes an ecological signal able to elicit robust activations throughout the auditory nervous system (see Pasley et al. 2012; Zion Golumbic et al. 2013; Leonard et al. 2015;

Nourski et al. 2015 for recent results on speech encoding within human auditory cortices), its complex spectro-temporal structure and its inherent communication purpose restricts a systematic manipulation of speech parameters. Here, we aimed at making a step forward in building a generalized model of the neural encoding of sound envelope by taking advantage of a complex and yet objectively quantifiable auditory stimulus: musical rhythms.

Entrainment to rhythms differs from speech and other behaviors at several levels: it is uniformly and widely expressed across all human societies, develops early in life, and has a long human evolution, thus making this behavior ideally suited for understanding non-verbal communication and its neural basis (see Hove and Risen 2009; Phillips-Silver and Keller 2012; Merchant and Honing 2014). Importantly, the structural parameters of a rhythm can be manipulated systematically (e.g., by gradual increase or decrease of the inter-onset intervals or by gradual deterioration of the pattern's regularity), thus facilitating modeling. Hence, musical rhythms appear to be a class of stimuli combining both complexity and flexibility of their structure. Because musical rhythms are usually more complex than mere isochronous sequences of tones, they are likely to induce more sophisticated temporal grouping of the acoustic events when listening to the incoming auditory stream. In particular, even when music is not strictly periodic, humans perceive periodic beats and spontaneously entrain body movements to this temporal periodic grid (e.g., London 2004; McAuley 2010; Phillips-Silver and Keller 2012 for some reviews). Hence, as described by music theorists, musical beat is not itself a stimulus property, although it is usually induced by a rhythmic stimulus.

The primary goal of the current study is to test whether the processing of a rhythmic pattern in the human primary auditory cortex is characterized by a neural transform of the acoustic input related to the spontaneous perception of a beat in this rhythm. To achieve this goal, we took advantage of a rare opportunity to record directly in the human auditory cortex from depth-electrodes implanted within the Heschl's gyrus and the planum temporale, i.e., two brain areas usually considered as the primary and secondary auditory cortex in humans (see Da Costa et al. 2011), in eight patients undergoing surgery for the treatment of intractable epilepsy (total of 55 recording sites) (Bancaud and Talairach 1973). Second, we identified the intracerebral EEG responses frequency-locked to the rhythm envelope in these areas by means of a frequency-tagging approach (Nozaradan 2014 for review). Using this approach, the previous scalp EEG studies have showed that frequencies of the EEG activity elicited by rhythmic inputs are amplified when they correspond to the perceived beat frequency, even when a sound does not occur on each beat,
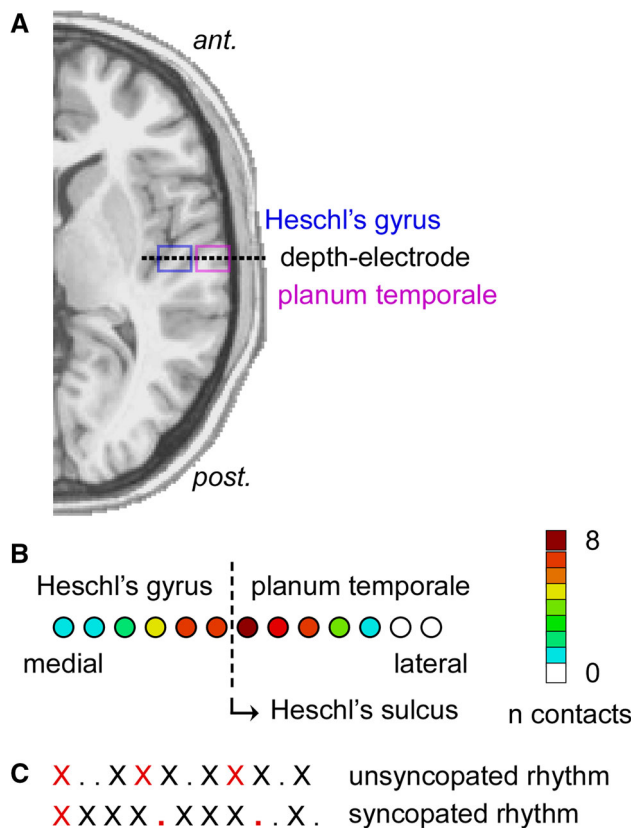
Fig. 1 Electrode implantation and rhythm stimuli. **a** Schematic representation of a depth-electrode (*dashed line*) implanted within Heschl's gyrus (contacts inside the *blue box*) and the planum temporale (contacts inside the *purple box*). **b** Number of contacts located in Heschl's gyrus and the planum temporale across all patients. Heschl' sulcus (*dotted line*) served as reference line to delimit the two gyri. **c** Schematic representation of the unsyncopated and syncopated auditory rhythms played to the participants (*X* 200 ms white noise burst; *dots* 200 ms silent intervals). The location of the beat is shown in *red*, based on the previous work (Nozaradan et al. 2016a). Note that in the unsyncopated rhythm, all beats coincide with the occurrence of a sound, whereas in the syncopated rhythm, some beats coincided with silent intervals

i.e., in syncopated rhythms (Nozaradan et al. 2011, 2012b; Chemin et al. 2014).

Here, the intracerebral EEG was recorded, while the patients listened to different rhythms. These rhythms were selected based on the previous work showing that they induce a consistent beat across individuals and a relative enhancement of the neural activity at this perceived beat frequency, as recorded with scalp surface EEG (Nozaradan et al. 2012b, 2016a). These rhythms consisted of white noise bursts alternating with silences (Fig. 1). The frequency spectrum of the envelope of these rhythms did not contain a single frequency, but multiple frequencies within the musical tempo range. Importantly, the different rhythms were composed of the same number of sounds and silent intervals, but differed with respect to the fact that one

rhythm was unsyncopated (i.e., a sound always occurred on the beat), whereas the other was syncopated (i.e., a sound did not occur on each beat) (Nozaradan et al. 2016a) (Fig. 1).

Our main hypothesis was that both Heschl's gyrus and the planum temporale would enhance frequency components coinciding with the beat, indicating that the widespread propensity to entrain to the beat when listening to musical rhythms could involve the earliest stages of auditory cortical processing. This selective enhancement observed in the primary cortical auditory areas could be the product of non-linear processing arising along the ascending auditory pathway before the cortex in response to sounds. This hypothesis was derived from a line of research investigating the extent to which musical conventions and habits may be determined by evolutionarily shaped human physiology (see Hove et al. 2014; Rajendran et al. 2015). This selective enhancement could also be due to top-down modulation of the neural response in these auditory cortices by higher level associative areas and/or motor areas (Patel and Iversen 2014; Large et al. 2015). An alternative hypothesis was that selective gain at beat frequency occurs only in the planum temporale. Since the planum temporale appears to play a key role in higher level representation of incoming sounds (Griffiths and Warren 2002), this would indicate that this characteristic input–output transform as observed with scalp recordings requires higher level perceptual organization. Finally, if these auditory areas were not found to transform the input towards a selective enhancement of activities at beat-related frequencies, this would indicate that the selective gain observed with scalp recordings critically involves higher level associative areas and/or motor areas with no direct impact on the neural response to rhythms in primary auditory cortices.

Importantly, here, we did not restrict our analyses to the frequency content of the low-frequency components of intracerebral EEG signals. Indeed, we also assessed the envelope of high-frequency intracerebral EEG activities (>30 Hz) (Brosch et al. 2002; Brugge et al. 2009; Nourski et al. 2009; Gourévitch et al. 2011; Steinschneider et al. 2011, 2013). Whereas low-frequency activities have been proposed to mainly reflect synaptic activity, high-frequency activities have been proposed to relate to spiking activity (Steinschneider et al. 2008; Miller et al. 2012). Interestingly, recent studies using speech stimuli have revealed differences related to stimulus acoustics and perception between the low- and high-frequency responses in Heschl's gyrus and the planum temporale (e.g., Nourski et al. 2009, 2015), thus raising the possibility that low- and high-frequency activities reflect distinct aspects of sound encoding.

Materials and methods

Patients

Eight patients with medically intractable epilepsy (1 male, 7 females, all right-handed, age 29 ± 10 years, mean $\pm$ standard deviation) took part in the study after providing written informed consent. They had no musical experience in Western music more than as listeners or occasional dancers and no prior experience with the task used in this study. Their intellectual performances were in a normal range and compatible with the understanding and realization of the task. There was no history of hearing disorder. None of the patients have auditory auras during their epileptic seizures. SEEG explorations showed anterior temporal epilepsy in five patients, basal temporal epilepsy in two patients, and insular epilepsy in one patient. None of these patients showed a seizure onset zone in the auditory cortex. One patient has a history of depression. The study was approved by the local ethics committee (CHU Nancy, France) and carried out in accordance with the approved guidelines, in agreement with the declaration of Helsinki.

Stereo-encephalographic (SEEG) placement of intracerebral electrodes

SEEG was performed according to a standard clinical protocol to localize and delineate the areas of epileptic seizure onset and early propagation and, thereby, plan a surgical treatment of epilepsy (see Bancaud and Talairach 1973; Jonas et al. 2016). Electrode implantation sites were chosen according to non-invasive data collected during the earlier phase of the investigation. The intracerebral electrode rods consisted of 5–18 contiguous 2 mm contacts separated by 1.5 mm (Dixi Medical, Besançon, France). All electrodes targeted the posterior superior part of the insular cortex, crossing over Heschl's gyrus and the planum temporale, respectively (Fig. 1; see Table 1 for the MNI coordinates of the contacts of interest).

Anatomical localization of electrode contacts

A total of 55 contacts of interest recording local field potentials were considered in the current study. Electrode contacts in Heschl's gyrus and the planum temporale were identified using the individual MRI data acquired before the implantation merged with the CT scan imagery acquired just after the implantation (Iplanstereotaxy; BrainLab, Feldkirchen, Germany) (Jonas et al. 2016). Heschl's gyrus and the planum temporale were delimited based on anatomy, with Heschl' sulcus as reference point (see Fig. 1). On average, three contacts were located in Heschl's gyrus (0–6 contacts) and three contacts were located in the planum temporale (2–5 contacts) (Fig. 1). The tissue in which these electrode contacts were located was overall free of epileptic activity.

Auditory stimuli

Auditory stimuli were created using Audacity 1.2.6 (http://audacity.sourceforge.net/) and presented binaurally through insert earphones at a comfortable hearing level (around 70 dB) using Matlab 2012 (The MathWork, USA). The stimuli consisted of two different rhythmic patterns (Fig. 1). They lasted 2.4 s and were looped continuously during 40 s. The rhythms were based on an alternation of 200-ms white noise bursts (linear 12-ms rise and 50 ms fall times) and 200-ms silences. White noise bursts were used to optimize the emergence of a response to the sound envelope independently of any tonotopic organization that could have been observed in the areas of interest (Edwards and Chang 2013). The rhythms used in the current study were selected based on the previous evidence that they induce a beat at a consistent frequency across individuals and that they elicit in the scalp surface EEG a relative enhancement of the neural activity at this beat frequency correlating with behavioral measures of sensorimotor synchronization to this perceived beat (Nozaradan et al. 2012b, 2016a). Specifically, these rhythms are known to

Table 1 Mean MNI coordinates for the contacts located in the Heschl's gyrus and the planum temporale for each patient

	Heschl's gyrus				Planum temporale			
	x	y	z	Number of contacts	x	y	z	Number of contact
Patient 1	−45	7	−6	3	−53	8	−6	2
Patient 2				0	49	−12	4	4
Patient 3	−40	−28	11	6	−57	−22	11	5
Patient 4	44	−12	7	3	49	−13	7	5
Patient 5	45	−16	−1	2	57	−14	3	5
Patient 6	45	−26	6	4	58	−26	10	4
Patient 7	−41	−25	1	2	−50	−25	1	3
Patient 8	−36	−26	7	4	−48	−27	7	3

induce the perception of a beat based on a grouping by four elements, at points represented in Fig. 1 (Nozaradan et al. 2012b, 2016a; Povel and Essens 1985). The first rhythm was *unsyncopated*, as a sound always occured on the beat. The second rhythm was *syncopated*, as a sound did not systematically occur on each beat (Fig. 1).

To test whether the hypothesized selective increase of the neural response at beat frequency was specific to a frequency range corresponding to the ecological range at which musical beats are perceived (<5 Hz; see van Noorden and Moelants 1999), patients were also presented with the unsyncopated rhythm played at four times the original tempo (unsyncopated rhythm tempo × 4) (Nozaradan et al. 2012b). This faster tempo was obtained by reducing the duration of the sounds and silent intervals from 200 ms to 50 ms, thus corresponding to 20 Hz (with proportional adaptation of the linear rise and fall time). At this speeded tempo, the perceived beat should no longer coincide with a grouping by four elements, as such a grouping would correspond to a beat frequency of 5 Hz (4×50 ms = 200 ms, or 300 bpm), which is too fast to be spontaneously perceived as a beat. Instead, the perceived beat coincides with an integer multiple grouping of elements, namely, a grouping by twelve elements, thus corresponding to a beat frequency of 1.66 Hz (Nozaradan et al. 2012b). Each rhythmic sequence of 40 s was presented four times. The order of presentation was counterbalanced across patients.

SEEG recording

SEEG was recorded continuously at a 512 Hz sampling rate using a 128-channel amplifier (2 SD LTM 64 Headbox, Micromed, Italy). The patients were comfortably seated in a chair with the head resting on a support. They were instructed to relax, avoid any unnecessary head or body movement, and keep their eyes fixated on a central dot in front of them to avoid any eye movements. A prefrontal-central surface electrode (Fpz) served as reference. The onset of each rhythm was triggered by the experimenter, when the patients felt ready to listen to the following trial. The patients were asked to focus the attention on the rhythms and to tell at the end of each sequence whether the sequence was identical or not from the preceding one (thus checking for a minimum attention allocated to the sequences throughout the listening). All patients responded with no errors.

Low-frequency (<30 Hz) responses

Analysis of SEEG data was carried out using Letswave 5 (http://nocions.webnode.com) running under Matlab 2012. The continuous SEEG files were filtered using a 0.1-Hz

high-pass FFT filter to remove very slow drifts in the recorded SEEG. Epochs lasting 40 s were obtained by segmenting the recordings from +0 to +40 s relative to the onset of each auditory sequence. For each patient and rhythm, the four SEEG epochs were averaged across trials to enhance the signal-to-noise ratio (SNR) of SEEG activities time-locked to the rhythms.

The obtained average waveforms were transformed in the frequency domain using a discrete Fourier transform, yielding a frequency spectrum of signal amplitude (μV) ranging from 0 to 256 Hz with a frequency resolution of 0.025 Hz. This procedure allowed SEEG responses frequency-locked to the frequency components of the sound envelope to be quantified, to objectively assess the neural transformation of each sequence in the brain areas of interest. We deliberately analyzed long-lasting epochs to maximize the frequency resolution of the Fourier transform. This ensures that the response of interest is captured in a narrow frequency band, thereby enhancing SNR, and making it possible to disentangle signals located at nearby frequencies (see Regan 1989; Rossion 2014; Norcia et al. 2015; see Jonas et al. 2016 for this procedure applied to SEEG recordings following periodic visual stimulation).

Within the obtained frequency spectra, SEEG amplitudes may be expected to correspond to the sum of (1) peaks elicited by the stimuli (i.e., the signal) and (2) unrelated background noise due, for example, to spontaneous SEEG activity. Therefore, to estimate the responses to the auditory rhythms, the contribution of background noise was removed by subtracting, at each bin of the spectra, the average amplitude measured at neighboring frequency bins (2 frequency bins ranging from −3 to −5 and from +3 to +5 bins of 0.025 Hz width). The validity of this subtraction procedure relies on the assumption that, in the absence of response, the amplitude at a given frequency bin should not differ from the mean amplitude of the surrounding frequency bins (Mouraux et al. 2011; Wang et al. 2012; Nozaradan et al. 2012a, 2015; Jonas et al. 2016). The magnitude of the signal at the frequencies corresponding to the sound envelope was then estimated by taking the maximum noise-subtracted amplitude measured in a range of three frequency bins centered on the expected response frequency of interest (i.e., based on the sound envelope spectrum) (Nozaradan et al. 2011, 2012b).

High-frequency (30–100 Hz) responses

To identify high-frequency responses, the waveforms were filtered using a 25-Hz high-pass FFT filter to prevent contamination of the time–frequency transforms by low-frequency responses elicited by the sound envelope beneath 25 Hz. This high-pass cut-off was chosen to consider the frequency of the individual sounds in the case of the speeded

rhythm (50-ms duration, i.e., 20 Hz), thus keeping the analysis comparable across the rhythms. A 50-Hz notch filter was also applied to minimize contamination with power-line noise. A time–frequency decomposition of the SEEG was then performed using the Morlet wavelet transform (time standard deviation of the wavelet: $\sigma = 8$). Estimated frequencies ranged from 30 to 100 Hz in 2-Hz increments. The modulus of the wavelet transform was used to express oscillation amplitude as a function of time and frequency.

To assess whether the envelope of these high-frequency responses contained frequencies corresponding to the frequencies of the sound envelope, the envelope of the high-frequency SEEG responses (averaged between 30 and 100 Hz) was processed using the same procedure as for the low-frequency activities (Fig. 2). Specifically, the high-frequency SEEG activities were transformed using a discrete Fourier transform, yielding a frequency spectrum of signal amplitude (μV) ranging from 0 to 256 Hz with a frequency resolution of 0.025 Hz. Then, noise-subtracted spectra were computed using the same spectral baseline subtraction procedure as for the low-frequency responses.

Sound analysis

To determine the frequencies at which the peaks of response could be expected in the intracerebral EEG spectra, the temporal envelope of the 40-s sound sequences

was extracted using a Hilbert transform as implemented in Matlab and transformed in the frequency domain using a discrete Fourier transform, yielding a frequency spectrum of acoustic energy (Nozaradan et al. 2016a, b; Cirelli et al. 2016). The envelope of the rhythms consisted of 12 distinct frequency components ranging from the frequency corresponding to the period of the entire rhythm to the frequency corresponding to the period of the individual elements at each tempo of presentation (Figs. 3, 4, 5a).

Statistical analysis

Statistical analyses were performed using SPSS Statistics 21.0 (IBM, Armonk, NY, USA). Significance level was set at $p < 0.05$.

For each patient, spectra were averaged across the electrode contacts located within Heschl's gyrus, and the planum temporale, respectively. For each rhythm, a one-sample t test was then used to determine whether the average noise-subtracted response amplitudes obtained within Heschl's gyrus and the planum temporale for the low- and high-frequency responses were significantly greater than zero. Indeed, in the absence of significant peaks, the averaged noise-subtracted peak amplitudes should tend towards zero.

Our hypothesis required to test whether there was an actual increase of the SEEG response at beat frequency, or whether the SEEG response was a mere tracking of the
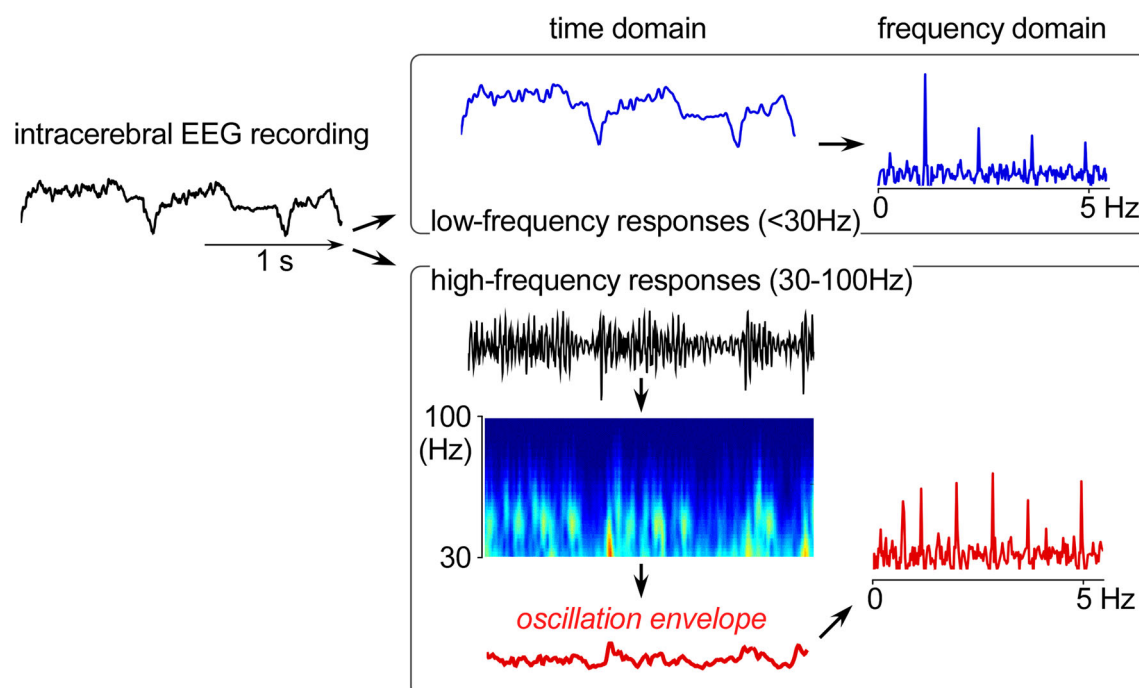

Fig. 2 Schematic representation of the frequency-tagging approach used to isolate low- and high-frequency intracerebral EEG responses related to processing the sound envelope. Low-frequency responses (*upper panel*) were obtained by Fourier transform of the EEG epochs. Modulation of high-frequency EEG responses (*bottom panel*) was obtained by extracting the envelope of broadband activities using a Morlet wavelet transform to estimate the amplitude of high-frequency activities between 30 and 100 Hz (time–frequency map) and by Fourier transform of the resulting envelope
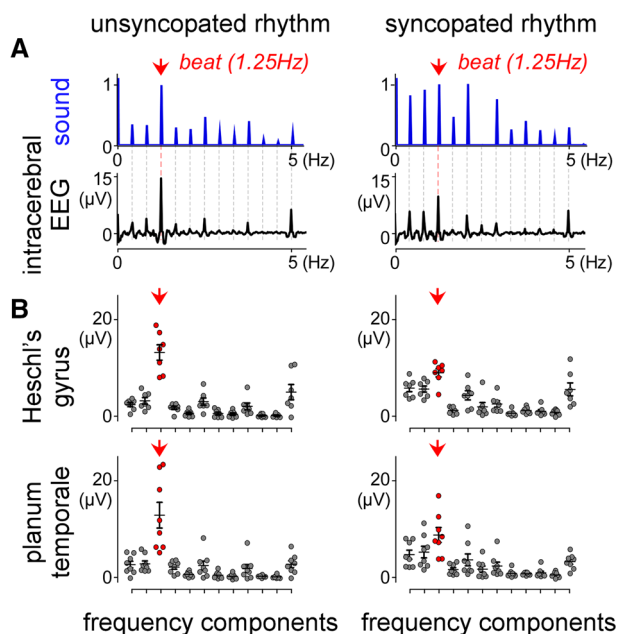
Fig. 3 Sharpening of the beat-related frequency in the auditory cortex. Low-frequency intracerebral EEG responses to the unsyncopated (*left*) and syncopated (*right*) rhythms. **a** Spectrum of the sound envelope of the rhythm sequences and the corresponding intracerebral EEG spectra averaged across all contacts located in Heschl's gyrus and the planum temporale, averaged across patients. The *red arrow* indicates the beat frequency. **b** Amplitude of all frequency components elicited by the rhythm sound envelope in Heschl's gyrus and the planum temporale. Each *dot* represents the amplitude, in microvolts, of the intracerebral EEG response averaged across contacts located in Heschl's gyrus or the planum temporale for each patient. The *whisker bars* represent the mean ± standard deviation across participants. Note the neural transform of the rhythms characterized by a relative increase of the response elicited at the frequency corresponding to the beat. This was the case not only for the unsyncopated rhythm, but also for the syncopated rhythm, in which other frequencies were almost as strong as the beat frequency in the envelope spectrum of the rhythm. Note also the strong response observed at 5 Hz providing evidence against a simple low-pass explanation below 5 Hz. One possible explanation for this observation could be that the frequency corresponding to the rate at which the individual sounds were presented constituted a relatively salient feature for the responding neural populations

Fig. 4 Lack of sharpening of the beat-related frequency for speeded rhythms (low-frequency intracerebral EEG responses). **a** Spectrum of the sound envelope of the rhythm sequences and the corresponding intracerebral EEG spectra averaged across all contacts located in Heschl's gyrus and the planum temporale, averaged across patients. **b** Amplitude of all frequency components elicited by the rhythm sound envelope in Heschl's gyrus and the planum temporale. Note that the selective increase of the intracerebral EEG response to the beat frequency when participants listened to the original version of the unsyncopated rhythm (*left* graphs, *red arrow*) was no longer present when the rhythm was played four times faster (*right* graphs, *light red arrow*). Instead, a slower frequency component corresponding to a grouping by 12 elements was selectively increased (*light red arrow*), suggesting a low-pass function in the input–output transform of rhythmic sequences in these auditory cortical areas

rhythmic input with no relative increase at any particular frequency. To test this hypothesis, an index of relative SEEG amplitude at beat frequency was calculated for each patient as the percentage of difference between the amplitude of the peak elicited at beat frequency minus the mean amplitude of the other peaks elicited by each rhythm. The same index was computed for the corresponding sound envelopes. Then, for each rhythm, this SEEG index was compared to the corresponding sound envelope index using a one-sample *t* test (Fig. 6). This direct SEEG-sound comparison specifically aimed to test the hypothesis of a relative enhancement or gain of the SEEG response to the beat, regardless of whether the beat coincided with a highly
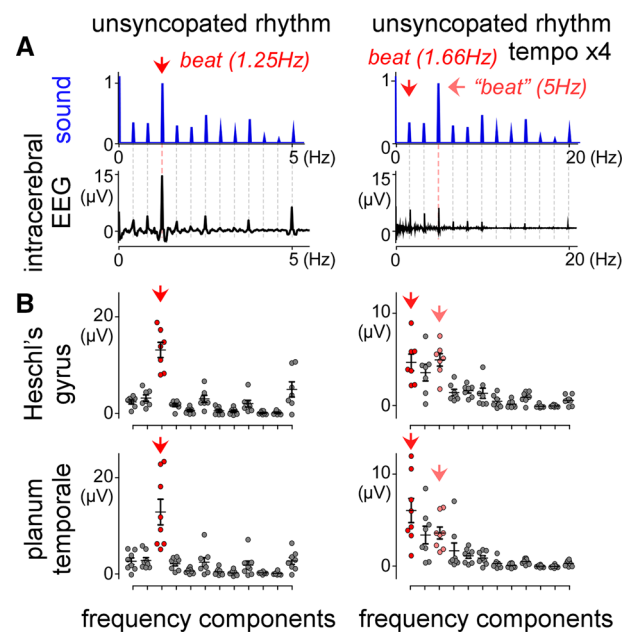
prominent frequency component of the sound envelope. In the syncopated rhythm, this sound-EEG comparison was also tested for two other frequency components which were almost as strong as the beat frequency in the sound envelope (frequency components 2 and 5, i.e., 0.832 and 2.083 Hz, respectively), thus testing the selectivity of the gain effect for the frequency component corresponding to the beat (Fig. 6, middle panel). In the speeded rhythm, this measure was conducted for the component corresponding to the prominent frequency in the sound envelope (frequency component 3, i.e., 5 Hz) but also for the slower frequency corresponding to a larger grouping of the events and assumed to correspond to the beat at that tempo (frequency 1, i.e., 1.66 Hz; see Nozaradan et al. 2012b) (Fig. 6, bottom panel).

In addition to these sound-EEG comparisons, this measure of relative amplitude was also compared across frequency bands (low- and high-frequency responses) and across the two locations (Heschl's gyrus and planum temporale) using two-way repeated-measures ANOVAs,
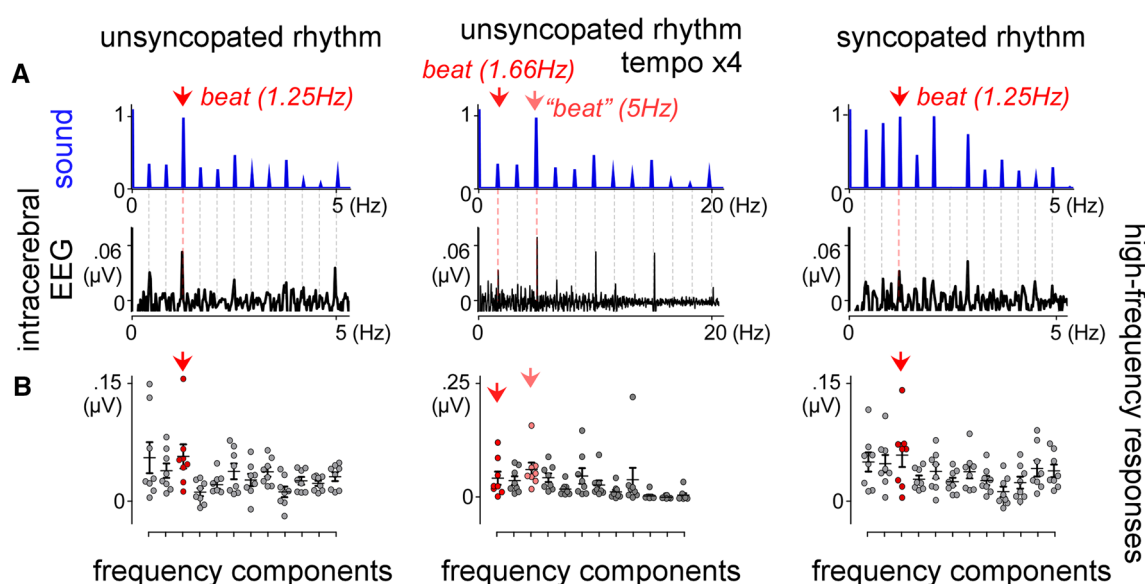
Fig. 5 High-frequency intracerebral EEG responses to the syncopated, unsyncopated, and speeded syncopated rhythms. **a** Spectrum of the sound envelope of the rhythm sequences and the corresponding intracerebral EEG spectra averaged across all contacts located in Heschl's gyrus and the planum temporale, averaged across patients. **b** Amplitude of all frequency components elicited by the rhythm sound envelope, averaged across Heschl's gyrus and the planum temporale. Note that, in contrast with the low-frequency responses, the input–output transform observed in the high-frequency responses (30–100 Hz) did not exhibit any selective increase of one particular frequency

thus allowing the relative amplitude of these responses to be compared across frequency bands and locations.

Phase analysis

The low- and high-frequency responses elicited by the rhythms in Heschl's gyrus and the planum temporale can be characterized by their amplitude, but also by their respective phase, to estimate the lag between these responses. These relative lags were obtained from the complex-valued Fourier transform of the epochs at the frequency component of 5 Hz, as it corresponds to the unitary acoustic elements in the rhythms (and to the unitary groups of elements at speeded tempo). Phase angles were converted to time lags according to the cycle length of 200 ms (Nozaradan et al. 2016b). These lag values were then compared using a two-way ANOVA with the factors 'Location' and 'Frequency band'. Hence, if significant, a difference would reflect a significant phase lag between frequency bands or locations, or a significant difference in the general tracking of the inputs.

Results

SEEG frequency components elicited by the auditory rhythms

As shown in Figs. 3, 4, and 5, each rhythmic stimulus elicited low- and high-frequency electrophysiological responses at frequencies corresponding to the sound envelope. Despite the short auditory stimulation time (less than 10 min in total), these responses were significantly greater than zero, and exhibited a very high SNR (see Table 2 for detailed values). These results support the previous evidence of sound envelope tracking in these two frequency bands in Heschl's gyrus and the planum temporale (see Brugge et al. 2009).

Sound-EEG comparison of low-frequency SEEG responses

Both in Heschl's gyrus and in the planum temporale, the measure of relative SEEG amplitude at beat frequency was significantly greater than the corresponding measure of relative amplitude obtained from the sound envelope. This was the case for both the *unsyncopated rhythm* and the *syncopated rhythm* (Fig. 6, red asterisks). This gain appeared to be selective to the frequency component coinciding with the beat, as other frequency components almost as strong as the beat frequency in the sound envelope of the syncopated rhythm were not significantly enhanced (i.e., frequencies 2 and 5 of the syncopated rhythm; Fig. 6, middle right panel). Moreover, this gain was disrupted at faster tempo (unsyncopated rhythm tempo x4), as the most prominent frequency component of the sound envelope in that rhythm did not elicit a significantly enhanced SEEG response (frequency 3, i.e., 5 Hz; Fig. 6, bottom panel). Instead, in that speeded rhythm, a significant increase of the SEEG response corresponding to a
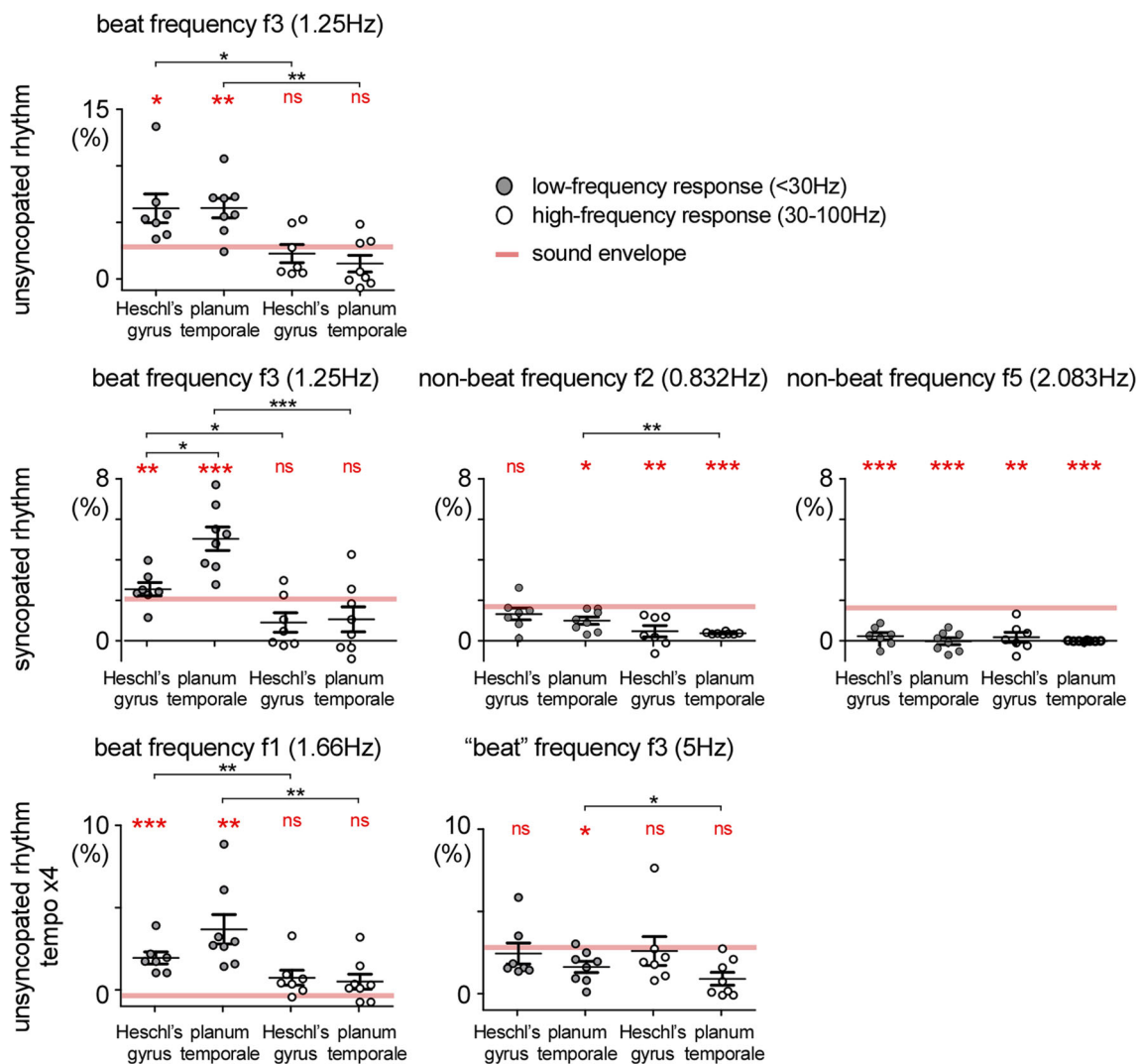
Fig. 6 Relative amplitude of the SEEG responses at beat frequency compared to the sound. Percentage of difference between the amplitude of the SEEG responses to specific frequency components of the sound envelope and the mean amplitude of the peaks at all other frequency components elicited by each rhythm. When participants listened to the unsyncopated and syncopated rhythms, the index of relative amplitude of the low-frequency responses to the beat (*grey dots*, mean ± standard deviation, each point corresponding to a patient) was significantly larger as compared to the corresponding amplitude of the beat frequency component in the sound envelope (*horizontal red line*). This relative increase of the beat frequency response was sharper in the planum temporale as compared to the Heschl's gyrus when participants listened to the more challenging

syncopated and fast conditions rhythms. This gain was selective to the beat frequency, as when participants listened to the syncopated rhythm, other frequencies of the sound envelope which were almost as strong as the beat frequency were not enhanced (*middle* graphs: frequencies f2 and f5). Moreover, when participants listened to the speeded unsyncopated rhythm (*bottom* graphs), a slower frequency (f1) was selectively enhanced, thus suggesting a low-pass function shaping this input–output transform. Contrasting with the low-frequency SEEG responses, the high-frequency SEEG responses (*white dots*) did not exhibit a significant increase at beat frequency. *Black asterisks*, p values obtained from post hoc paired-sample t tests. *Red asterisks*, p values obtained from the one-sample t tests against sound. *ns* non-significant; $*p < 0.05$; $**p \leq 0.01$; $***p \leq 0.001$

larger grouping by 12 elements was observed (frequency 1, i.e., 1.66 Hz; Fig. 6, bottom panel).

Sound-EEG comparison of high-frequency SEEG responses

Contrasting with the low-frequency SEEG responses, no significant input–output enhancement of the index of relative SEEG amplitude at beat frequency was observed in any of the

rhythms (Fig. 6, red asterisks), suggesting that low- and high-frequency SEEG responses reflect distinct processing of the sound inputs.

Relative amplitude at beat frequency in low- and high-frequency SEEG responses and in Heschl's gyrus and the planum temporale

For the *unsyncopated rhythm* (Fig. 6, upper panel), the repeated-measures ANOVA comparing the index of

Table 2 Frequency components elicited by the rhythms

	Unsyncopated rhythm	Syncopated rhythm	Unsyncopated rhythm tempo × 4
Low-frequency			
Heschl's gyrus	$t_6 = 6.77^{***}$	$t_6 = 7.26^{***}$	$t_6 = 8.14^{***}$
Planum temporale	$t_7 = 5.45^{***}$	$t_7 = 5.02^{**}$	$t_7 = 6.58^{**}$
High-frequency			
Heschl's gyrus	$t_6 = 3.18^{**}$	$t_6 = 3.41^{**}$	$t_6 = 4.58^{**}$
Planum temporale	$t_7 = 6.52^{***}$	$t_7 = 6.53^{***}$	$t_7 = 4.68^{**}$

On average, the rhythms elicited significant low- and high-frequency responses at the frequencies determined by the frequency components of the sounds envelope (one-sample t test against zero of the noise-subtracted amplitudes; ** $p \leq 0.01$; *** $p \leq 0.001$)

relative amplitude at beat frequency revealed a significant main effect of 'Frequency' in favor of the low-frequency responses ($F_{1,6} = 7.92$, $\eta^2 = 0.57$, $p = 0.03$), thus corroborating the evidence of an increased relative amplitude at beat frequency in the low-frequency responses but not the high-frequency responses. There was no significant effect of the factor 'Location' ($F_{1,6} = 0.31$, $\eta^2 = 0.05$, $p = 0.56$) and no interaction between the two factors ($F_{1,6} = 0.73$, $\eta^2 = 0.11$, $p = 0.42$). For the *syncopated rhythm* (Fig. 6, middle left panel), the ANOVA showed a significant main effect of the factor 'Frequency' (low-frequency responses > high-frequency responses, $F_{1,6} = 25.36$, $\eta^2 = 0.81$, $p = 0.002$), but also a main effect of the factor 'Location' (planum temporale responses > Heschl's gyrus responses, $F_{1,6} = 11.75$, $\eta^2 = 0.66$, $p = 0.01$). The marginally significant interaction between the two factors ($F_{1,6} = 6.15$, $\eta^2 = 0.50$, $p = 0.05$) was explained by the fact that the magnitude of low-frequency responses but not high-frequency responses were significantly greater in the planum temporale as compared to Heschl's gyrus (low-frequency responses: $p = 0.01$; high-frequency responses: $p = 0.52$). This indicates that in the case of the syncopated rhythm, i.e., when the beat was more ambiguous, the selective gain of low-frequency responses at beat frequency was sharper in the planum temporale as compared to Heschl's gyrus.

When the unsyncopated rhythm was presented at tempo x4, the strongest frequency component which corresponded to the beat at musical tempo (Fig. 6 bottom right panel, frequency component 3) showed no longer significant main effect of 'Frequency' ($F_{1,6} = 0.26$, $\eta^2 = 0.04$, $p = 0.62$), no main effect of 'Location' ($F_{1,6} = 3.32$, $\eta^2 = 0.35$, $p = 0.11$), and no interaction between the two factors ($F_{1,6} = 0.46$, $\eta^2 = 0.07$, $p = 0.520$), thus indicating a disruption of the gain effect at speeded tempo. However, when the ANOVA was performed using the relative amplitude of the SEEG response at the frequency corresponding to a grouping by 12 elements (frequency 1, Fig. 6, bottom left panel), there was a significant main effect of 'Frequency' ($F_{1,6} = 18.36$, $\eta^2 = 0.75$, $p = 0.005$)

and a significant main effect of 'Location' ($F_{1,6} = 7.65$, $\eta^2 = 0.56$, $p = 0.03$). Post hoc t tests revealed a tendency towards a greater increase of low-frequency responses in the planum temporale as compared to Heschl's gyrus for low-frequency responses ($p = 0.06$), but not for high-frequency responses ($p = 0.87$; see Fig. 6, black asterisks).

Phases lags

Low-frequency SEEG responses

For the unsyncopated rhythm, the mean latency of the low-frequency responses at the unitary element frequency was 121 ± 50 ms (mean and standard deviation) in Heschl's gyrus and 125 ± 40 ms in the planum temporale. For the syncopated rhythm, the mean latency was 126 ± 42 ms in Heschl's gyrus and 145 ± 46 ms in the planum temporale. For the unsyncopated rhythm presented at tempo x4, the mean latency was 126 ± 42 ms in Heschl's gyrus and 145 ± 46 ms in the planum temporale.

High-frequency SEEG responses

For the unsyncopated rhythm, the mean latency of the high-frequency responses at the unitary element frequency was 90.3 ± 65.8 ms in Heschl's gyrus and 144 ± 45 ms in the planum temporale. For the syncopated rhythm, the mean latency was 79 ± 26 ms in Heschl's gyrus and 112 ± 48 ms in the planum temporale. For the unsyncopated rhythm presented at tempo x4, the mean latency was 107 ± 44 ms in Heschl's gyrus and 112 ± 58 ms in the planum temporale.

For the unsyncopated rhythm, the 2×2 ANOVA comparing low- and high-frequency responses across locations revealed no significant main effect of the factor 'Frequency' ($F_{1,6} = 0.018$, $p = 0.899$, $\eta^2 = 0.003$), a marginal effect of the factor 'Location' in favor of a longer latency in the planum temporale ($F_{1,6} = 5.825$, $p = 0.052$, $\eta^2 = 0.493$), and no interaction between the two factors ($F_{1,6} = 4.019$, $p = 0.092$, $\eta^2 = 0.401$). For the syncopated

and speeded rhythms, there was no main effect of 'Frequency', no main effect of 'Location', and no interaction between the two factors (all p values >0.127).

Discussion

With intracerebral EEG recordings performed directly in the human auditory cortex, we examined whether the response to rhythms is characterized by a neural transform which could account for the emergence of a periodic beat percept when listening to these rhythms. To this aim, we tested the specific hypothesis that Heschl's gyrus and the planum temporale would exhibit an increased relative magnitude in response to the frequency component of the sound envelope corresponding to the beat. Both in Heschl's gyrus and in the planum temporale, tracking of the rhythms was observed in the form of low-frequency (<30 Hz) and high-frequency (>30 Hz) responses locked to the sound envelope. Supporting this hypothesis, the low-frequency responses were characterized by a selective gain in amplitude at beat frequency which was, however, not observed for the high-frequency responses. Moreover, this selective gain of low-frequency responses to the beat was relatively sharper in the planum temporale, especially for the more ambiguous rhythm where the beat did not always coincide with a sound and where the frequency of the beat did not correspond to the highest component of the sound envelope. Even though part of this neural transform of rhythms could already take place in the early subcortical auditory processes, these observations indicate that the neural representation of rhythmic patterns is also shaped at the level of auditory cortices. Finally, this selective gain appeared to be disrupted when the rhythm was played at a fast tempo lying outside the musical tempo range.

Selective gain at beat frequency

The increased low-frequency neural response at beat frequency reflects a non-linear response characteristic of the auditory system. This may reflect the auditory system's selective enhancement of the strongest frequencies of the sound envelope. This could be true in the unsyncopated rhythm, in which the frequency component corresponding to the beat was the most prominent component. However, this was not the case in the syncopated rhythm, in which other frequency components were almost as strong as the beat frequency component, and yet, these components were not enhanced. Furthermore, when participants listened to the speeded version of the unsyncopated rhythm, there was no selective gain of the highest frequency component of the sound envelope. Instead, there was a selective gain of a lower frequency component below 5 Hz, thus lying within

the ecological range for beat perception and corresponding to the most-often perceived beat at that tempo (see Nozaradan et al. 2012b), corresponding to a grouping by 12 elements. These observations indicate that the non-linear transform of the sound input into neural activity in these auditory areas is not merely the consequence of an enhancement of the strongest components of the sound envelope, but also involves other mechanisms.

Interestingly, this selective gain at beat frequency has previously been predicted by models of non-linear resonance (Large 2010; Velasco and Large 2011). According to these models, the emergence of a selective entrainment at beat frequency could be explained as the non-linear product of the coupling between the rhythmic input and neural oscillators (the term "entrainment" referring here to a synchronization process by which the frequency components of the output signal correspond to those of the input; see Large et al. 2015). The non-linear characteristics of these oscillators could thus explain the emergence of a beat percept when listening to complex rhythms in which the beat frequency has no prominent acoustic energy in the rhythmic input, especially in syncopated rhythms in which the perceived beat does not systematically coincide with an acoustic event (see Large 2010; Velasco and Large 2011 for more details on the model). Hence, our current findings could reflect, at least in part, the product of this non-linear dynamic system as embodied in the auditory system. Overall, these mechanisms could account for the observation that, while many frequency and phase combinations are available in a musical rhythm, different individuals listening to this kind of rhythm tend to select a similar beat frequency, because this frequency is made salient by the responding system.

Low-pass frequency constraint

Beat perception is known to occur within a specific range of frequencies corresponding to the musical tempo (∼0.5–5 Hz, i.e. 30–300 beats per min or bpm) (van Noorden and Moelants 1999). This frequency range is explained in part by the biomechanical properties of the human body constraining the execution of repeated body movements in synchronization to sounds, and consequently constraining beat perception itself (McAuley 2010; Toivi-ainen et al. 2010). However, this frequency range might also be due to timing constraints inherent to the neural network responding to the stream of rhythmic auditory inputs (Large 2008; Lakatos et al. 2008; Schroeder and Lakatos 2009).

Here, we found that speeding up the tempo of the rhythm, such that the original beat frequency lies outside the musical tempo range for beat perception disrupted this characteristic neural transform. That is, the peak elicited at "beat" frequency was no longer selectively enhanced. In

contrast, the peak elicited at slower frequency corresponding to a larger perceptual grouping by 12 elements (instead of 4 elements) became selectively enhanced at the fast tempo. This finding provides evidence for sensitivity of the neural transform of rhythms in favor of a frequency range lower than ~ 5 Hz. Such a frequency tuning function shaping the neural responses to rhythms in the primary auditory cortices could explain the range of frequencies in which the beat is usually perceived and produced in music (van Noorden and Moelants 1999; Large 2008; Nozaradan et al. 2012b).

This result can also be related to the previous observations of a band-pass tuning of the sensitivity to modulated sounds in the ~ 2–5-Hz range characterizing human auditory perception (Edwards and Chang 2013). In acoustic studies, this sensitivity was observed in response to different types of stimuli (e.g., periodic amplitude- or frequency-modulated tones or narrow-band noise; see Edwards and Chang 2013 for a review). Moreover, this range of frequency modulation has been shown to constitute a key component of the frequency content of the sound envelope of speech, given that it conveys critical information for speech intelligibility (Drullman et al. 1994a, b; Shannon et al. 1995). The current results thus add to this line of research by showing that while the human primary auditory cortices present the capacity to frequency-lock to sound envelopes over a wide range of frequencies beyond 5 Hz, the characteristic neural transform of rhythms that we observed in these areas is shaped by a low-pass function corresponding to the low-pass function observed in behavior.

Spatial distribution

Responses frequency-locked to the rhythm envelope were obtained in both Heschl's gyrus and the planum temporale, i.e., cortical regions that are assumed to correspond to the primary and secondary auditory cortex, respectively. There was no significant difference in the relative phase of the responses recorded from the two regions, suggesting the possibility of a parallel processing of sound inputs in the two areas.

Interestingly, the selective gain at beat frequency was relatively sharper in the planum temporale as compared to Heschl's gyrus, especially for the rhythms in which the beat did not correspond to the most prominent component of the sound envelope (i.e., the syncopated rhythm, in which the beat frequency was not as prominent as in the unsyncopated rhythm, and the speeded rhythm, in which the selective gain was not observed at the most prominent peak of the sound envelope spectrum). Since the planum temporale is thought to play a key role in higher level representation of incoming sounds than Heschl's gyrus

(Griffiths and Warren 2002), this observation suggests that this characteristic neural transform of rhythmic inputs involves higher level perceptual organization. The finding of a sharpened representation of beat-related responses in the planum temporale fits well with models of this region as a computational engine segregating the components of the incoming acoustic stream and matching these components with internal representations (Griffiths and Warren 2002). The temporal information derived from this matching would thus be gated to higher order cortical areas, eventually leading to higher order perceptual categorization or overt motor commands.

Whether this processing of rhythms in the auditory cortex requires connections with a network of distant brain areas remains unknown. The selective gain observed here in the auditory cortex could already arise in subcortical sites of the ascending auditory pathway: subcortical low-level processes of the auditory system combining band-pass constraints and mechanisms preferentially boosting prominent components of the sound envelope could result in a biasing effect on the perception of rhythms that might contribute to beat perception. Relatedly, recent models have proposed to quantify the extent to which beat perception can actually be explained by early auditory sensory processing in cochlear and brainstem auditory nuclei (Rajendran et al. 2015). The results of our study thus add to this line of research by providing the first description of the actual response to complex rhythms at the earliest cortical level of human auditory processing.

Alternatively, the characteristic neural transform of rhythms that we observed here in the auditory cortex may actually require connections between the ascending auditory pathway and other brain structures, in particular, motor areas, such as the premotor cortex, the basal ganglia, the cerebellum, and the supplementary motor area (Patel and Iversen 2014). This alternative hypothesis is based on the evidence of an automatic activation of these sensorimotor areas when perceiving a beat from auditory rhythms, even without overt movements, as captured with functional neuroimaging (see Zatorre et al. 2007; Grahn 2012 for reviews). Such functional connections between these regions could explain how information emerging from movement or conveyed by other sensory modalities, such as proprioception, may contribute to the internal representation of beat, by shaping this neural response (Phillips-Silver and Trainor 2007; Chemin et al. 2014). These functional connections could also account for the potential modulation of these neural transforms of rhythms according to top-down manipulations of the beat percept (as shown in the previous work, for example, by asking participants to voluntarily impose a mental imagery of meter on the sound input or by priming the beat representation via body movement; Nozaradan et al. 2011; Chemin et al.

2014). Hence, future research capturing the input–output transform of rhythms across different subcortical or cortical regions (via multiple depth-electrodes recordings or simultaneous brainstem-cortical recordings with scalp EEG; see Nozaradan et al. 2016b) and across different top-down manipulations of beat percept might provide a more extensive picture of these processes as embodied throughout the nervous system.

Low- vs. high-frequency responses

Both in Heschl's gyrus and the planum temporale, tracking of the rhythm envelope was observed in the form of both low-frequency and high-frequency responses. Low-frequency responses corresponded to variations in the recorded signal tracking directly the sound envelope, whereas high-frequency responses corresponded to modulations of the amplitude of high-frequency activities. Our result is in line with the previous observations showing that both areas respond to sound envelopes in these two frequency bands (Brugge et al. 2009; Nourski et al. 2009; Gourévitch et al. 2011; Miller et al. 2012; Steinschneider et al. 2011, 2013).

Interestingly, the spectra of the low-frequency responses were markedly similar to those obtained from human scalp recordings in healthy participants, characterized by an input–output transform in favor of frequency components corresponding to the beat (Nozaradan et al. 2012b, 2016a). This result also corroborates the previous observations of a tight correspondence between auditory evoked responses captured from the scalp and auditory evoked local field potentials recorded from primary auditory cortices (Nourski et al. 2015).

In contrast, no selective gain of the responses elicited at beat frequency was observed in the high-frequency neural activity (>30 Hz). This observation can be interpreted in light of the previous work showing that low-frequency responses, such as local field potentials, primarily represent synaptic rather than neuronal spiking activity, whereas high-frequency activities predominantly reflect spiking activity from neuronal aggregates (Friedman-Hill et al. 2000; Frien et al. 2000; Brosch et al. 2002; Miller et al. 2012; Nourski et al. 2009, 2015). Based on this assumption, the high-frequency responses that we isolated in this study could reflect processes produced by the mass firing of neuronal multiunits at a local scale. In contrast, the low-frequency responses that we identified could index the processing of widespread interconnected neural populations, thus reflecting processes that could be more closely related to perception (Schroeder and Lakatos 2009; Miller et al. 2012). Hence, low-frequency and high-frequency activities are likely to play different roles in sound encoding, as they convey different neural representations within the cortex.

Frequency-tagging to characterize the neural transforms of rhythms

Our study relied on the frequency-tagging approach to simultaneously measure low- and high-frequency responses frequency-locked to the sound envelope of rhythms with intracerebral EEG. This approach originally consists in periodically modulating a stimulus feature at a known frequency rate. This periodic modulation is expected to generate periodic activity within the neural populations responding to that stimulus feature, appearing as peaks in the EEG frequency spectrum at the corresponding frequency and/or its harmonics (Regan 1989; Norcia et al. 2015; see Jonas et al. 2016 for an application of this approach in human intracerebral recordings).

In the current study, the stimulation was not periodic but rhythmic, and the frequency-tagged responses were identified based on the frequency structure of the stimulus itself, i.e., the frequency components of the sound envelope. Even though each rhythmic sequence was repeated only four times, a high SNR was obtained for both the low- and high-frequency responses. This is due to the use of long-lasting sequences, providing a high spectral resolution (0.02 Hz) concentrating the frequency-locked activity sampled over a long-lasting period within narrow frequency bands of the EEG spectrum (Regan 1989; Rossion 2014). In addition, the use of long-lasting sequences is likely to improve the opportunity to capture dynamic aspects of the mechanisms at stake in the processing of rhythmic inputs, in contrast with transient event-related potentials (ERPs) reflecting transient neural responses triggered by the occurrence of transient stimuli (Nozaradan 2014).

Furthermore, these frequency-tagged activities can be directly compared with the corresponding frequency components of the auditory input (here the sound envelope) to characterize the neural transform of the acoustic input, thus sharing similarities with concepts such as the temporal modulation transfer function proposed to directly compare acoustic input with neural output (see Edwards and Chang 2013 for a review on this concept in the context of sound envelope). Here, this input–output comparison was necessary to show that there was an actual increase, i.e., a gain, of the EEG response at beat frequency regardless of whether this frequency coincided with a prominent frequency component of the sound envelope.

Functional significance, limitations, and perspectives

In this study, we did not gather explicit behavioral indices of beat perception. Therefore, we cannot ascertain that, in the explored patients, the rhythmic stimuli elicited a beat percept at the expected frequencies. However, we note that

there is currently no consensus on an unequivocal test by which beat perception could be behaviorally estimated (Tranchant and Vuvan 2015). Nevertheless, rhythmic patterns such as those of the current study have been used extensively in the previous work showing that, in healthy human individuals, they elicit the perception of a beat at a consistent frequency across individuals, as measured using behavioral tapping tasks (Nozaradan et al. 2012b, 2016a). Most importantly, frequency-tagging measures of neural activity at beat frequency have recently been found to be sensitive to individual differences related to sensorimotor synchronization and temporal prediction performance in tapping the beat on rhythmic sequences (Nozaradan et al. 2016a). The evidence of a correlation between EEG measures at beat frequency when listening to these rhythms on the one hand and behavioral measures of motor entrainment to the beat on these rhythms on the other hand corroborates the view that these neural responses might be somehow related to beat processing, thus providing indications on the functional significance of these neural measures.

Here, the current study reports a first description of the neural responses to these rhythms as directly recorded in the human auditory cortex. These responses were characterized by a non-linear transform of the sound envelope of the rhythms enhancing the beat frequency in these areas. These observations suggest that response properties from the earliest stages of auditory cortical processing may shape the neural representation of rhythmic inputs in favor of the emergence of a periodic beat percept. While beat perception might rely in part on these early cortical auditory processes, we do not claim that these processes fully account for the emergence of a beat percept and thus constitute a neural signature of beat perception. That is, a selective gain at beat frequency in these auditory cortical areas may constitute a necessary stage of beat perception process, which may nevertheless not be sufficient for an individual to actually perceive a beat. Hence, this study calls for future research to clarify the mechanisms by which these neural transforms arise in the different regions responding to rhythms (in the ascending auditory pathway and beyond) and how these transforms are linked to actual beat perception. For instance, this could be achieved by testing more conditions (including more rhythms and tempi) and by confronting the neural transforms of rhythms gathered across different regions with behavior and computational models, to understand the dynamics and functions of these neural transforms.

References

Bancaud J, Talairach J (1973) Methodology of stereo EEG exploration and surgical intervention in epilepsy. Rev Otoneuroophtalmol 45(4):315–328

Brosch M, Budinger E, Scheich H (2002) Stimulus-related gamma oscillations in primate auditory cortex. J Neurophysiol 87:2715–2725

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA 3rd (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol 102(4):2358–2374

Chemin B, Mouraux A, Nozaradan S (2014) Body movement selectively shapes the neural representation of musical rhythm. Psychol Sci 25(12):2147–2159

Cirelli LK, Spinelli C, Nozaradan S, Trainor LJ (2016) Measuring neural entrainment to beat and meter in infants: effects of music background. Front Neurosci 10:229. doi:10.3389/fnins.2016.00229

Da Costa S, van der Zwaag W, Marques JP, Frackowiak RS, Clarke S, Saenz M (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. J Neurosci 31(40):14067–14075

Drake C, Botte MC (1993) Tempo sensitivity in auditory sequences: evidence for a multiple-look model. Percept Psychophys 54(3):277–286

Drullman R, Festen JM, Plomp R (1994a) Effect of reducing slow temporal modulations on speech reception. J Acoust Soc Am 95:2670–2680

Drullman R, Festen JM, Plomp R (1994b) Effect of temporal envelope smearing on speech reception. J Acoust Soc Am 95:1053–1064

Edwards E, Chang EF (2013) Syllabic ($\sim$2–5 Hz) and fluctuation ($\sim$1–10 Hz) ranges in speech and auditory processing. Hear Res 305:113–134. doi:10.1016/j.heares.2013.08.017

Eggermont JJ (2001) Between sound and perception: reviewing the search for a neural code. Hear Res 157(1–2):1–42

Erulkar SD, Butler RA, Gerstein GL (1968) Excitation and inhibition in cochlear nucleus. II. Frequency modulated tones. J Neurophysiol 31:537–548

Fernald RD, Gerstein GL (1972) Response of cat cochlear nucleus neurons to frequency and amplitude modulated tones. Brain Res 45:417–435

Fraisse P (1967) Psychologie du temps. Presses universitaires de France, France

Friedman-Hill S, Maldonado PE, Gray CM (2000) Dynamics of striate cortical activity in the alert macaque: I. Incidence and stimulus dependence of gammaband neuronal oscillations. Cereb Cortex 10:1105–1116

Frien A, Eckhorn R, Bauer R, Woelbern T, Gabriel A (2000) Fast oscillations display sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey. Eur J Neurosci 12:1453–1465

Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. J Neurosci 32(5):1791–1802

Gourévitch B, Le Bouquin Jeannès R, Faucon G, Liégeois-Chauvel C (2011) Temporal envelope processing in the human auditory cortex: response and interconnections of auditory cortical areas. Hear Res 237(1–2):1–18

Grahn JA (2012) Neural mechanisms of rhythm perception: current findings and future perspectives. Top Cogn Sci 4(4):585–606

Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. Trends Neurosci 25(7):348–353

Hove MJ, Risen JL (2009) It's all in the timing: interpersonal synchrony increases affiliation. Soc Cogn 27:949–961

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 USA 111(28):10383–10388

Jonas J, Jacques C, Liu-Shuang J, Brissart H, Colnat-Coulbois S, Maillard L, Rossion B (2016) A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. Proc Natl Acad Sci USA 113(28):E4088–E4097

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84(2):541–577 (Review)

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320(5872):110–113

Large EW (2008) Resonating to musical rhythm: theory and experiment. In: Grondin Simon (ed) The psychology of time. Emerald, West Yorkshire

Large EW (2010) Neurodynamics of music. In: Riess Jones M, Fay RR, Popper AN (eds) Springer handbook of auditory research, vol 36., Music perceptionSpringer, New York, pp 201–231

Large EW, Herrera JA, Velasco MJ (2015) Neural networks for beat perception in musical rhythm. Front Syst Neurosci 25(9):159. doi:10.3389/fnsys.2015.00159

Leonard MK, Bouchard KE, Tang C, Chang EF (2015) Dynamic encoding of speech sequence probability in human temporal cortex. J Neurosci 35(18):7203–7214

Liégeois-Chauvel C, Lorenzi C, Trébuchon A, Régis J, Chauvel P (2004) Temporal envelope processing in the human left and right auditory cortices. Cereb Cortex 14(7):731–740

London J (2004) Hearing in time: psychological aspects of musical meter. Oxford UP, London

Malone BJ, Schreiner CE (2010) Time-varying sounds: amplitude envelope modulations. In: Rees A, Palmer AR (eds) The auditory brain. Oxford University Press, Oxford, New York, pp 125–148

McAuley JD (2010) Tempo and rhythm. In Jones MR et al. (eds.) Music Perception, Springer Handbook of Auditory Research 36, USA

Merchant H, Honing H (2014) Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. Front Neurosci 7:274. doi:10.3389/fnins.2013.00274

Miller KJ, Foster BL, Honey CJ (2012) Does rhythmic entrainment represent a generalized mechanism for organizing computation in the brain? Front Comput Neurosci 6:85

Møller AR (1972) Coding of amplitude and frequency modulated sounds in the cochlear nucleus of the rat. Acta Physiol Scand 86:223–238

Mouraux A, Iannetti GD, Colon E, Nozaradan S, Legrain V, Plaghki L (2011) Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. J Neurosci 31:6079–6087

Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B (2015) The steady-state visual evoked potential in vision research: a review. J Vis 15(6):4

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29(49):15564–15574

Nourski KV, Steinschneider M, Rhone AE, Oya H, Kawasaki H, Howard MA 3rd, McMurray B (2015) Sound identification in human auditory cortex: differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. Brain Lang 148:37–50

Nozaradan S (2014) Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. Philos Trans B 369(1658):20130393. doi:10.1098/rstb.2013.0393

Nozaradan S, Peretz I, Missal M, Mouraux A (2011) Tagging the neuronal entrainment to beat and meter. J Neurosci 31:10234–10240

Nozaradan S, Peretz I, Mouraux A (2012a) Steady-state evoked potentials as an index of multisensory temporal binding. Neuroimage 60(1):21–28

Nozaradan S, Peretz I, Mouraux A (2012b) Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. J Neurosci 32(49):17572–17581

Nozaradan S, Zerouali Y, Peretz I, Mouraux A (2015) Capturing with EEG the neural entrainment and coupling underlying sensori-motor synchronization to the beat. Cereb Cortex 25(3):736–747

Nozaradan S, Peretz I, Keller PE (2016a) Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. Sci Rep 6:20612. doi:10.1038/srep20612

Nozaradan S, Schönwiesner M, Caron-Desrochers L, Lehmann A (2016b) Enhanced brainstem and cortical encoding of sound during synchronized movement. Neuroimage 16:30322–30326

Pantev C, Hoke M, Lehnertz K, Lütkenhöner B, Anogianakis G, Wittkowski W (1988) Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalogr Clin Neurophysiol 69(2):160–170

Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight RT, Chang EF (2012) Reconstructing speech from human auditory cortex. PLoS Biol 10:e1001251

Patel AD, Iversen JR (2014) The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Front Psychol 8:57. doi:10.3389/fnsys.2014.00057

Phillips-Silver J, Keller PE (2012) Searching for roots of entrainment and joint action in early musical interactions. Front Hum Neurosci 6:26

Phillips-Silver J, Trainor LJ (2007) Hearing what the body feels: auditory encoding of rhythmic movement. Cognition 105(3):533–546

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 82:165–178

Povel DJ, Essens PJ (1985) Perception of temporal patterns. Music Percept 2:411–441

Rajendran VG, Harper NS, Willmore BD, Schnupp JWH (2015) A biologically plausible model of beat detection in complex rhythmic sounds. In: Proceedings of the rhythm perception and production workshop, Amsterdam

Regan DM (1989) Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. Elsevier, New York

Repp BH (2005) Sensorimotor synchronization: a review of the tapping literature. Psychon Bull Rev 12(6):969–992

Repp BH, Su YH (2013) Sensorimotor synchronization: a review of recent research (2006–2012). Psychon Bull Rev 20(3):403–452. doi:10.3758/s13423-012-0371-2

Rossion B (2014) Understanding individual face discrimination by means of fast periodic visual stimulation. Exp Brain Res 232(6):1599–1621

Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci 32(1):9–18

Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303–304

Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. Nature 416(6876):87–90

Steinschneider M, Fishman YI, Arezzo JC (2008) Spectrotemporal analysis of evoked and induced electroencephalographic

responses in primary auditory cortex (A1) of the awake monkey. Cereb Cortex 18(3):610–625

Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA 3rd (2011) Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. Cereb Cortex 21(10):2332–2347

Steinschneider M, Nourski KV, Fishman YI (2013) Representation of speech in human auditory cortex: is it special? Hear Res 305:57–73

Toiviainen P, Luck G, Thompson M (2010) Embodied meter: hierarchical eigenmodes in music-induced movement. Music Percept 28:59–70

Tranchant P, Vuvan D (2015) Current conceptual challenges in the study of rhythm processing deficits. Front Neurosci 9:197. doi:10.3389/fnins.2015.00197

van Noorden L, Moelants D (1999) Resonance in the perception of musical pulse. J New Music Res 28:43–66

Velasco MJ, Large EW (2011) Pulse detection in syncopating rhythms using neural oscillators. In: Proceedings of the 12th annual conference of the international society for music information retrieval, pp 186–190

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):2033–2041. doi:10.1152/jn.00310.2011

Zatorre RJ, Chen JL, Penhune VB (2007) When the brain plays music: auditory-motor interactions in music perception and production. Nat Rev Neurosci 8(7):547–558

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(5):980–991