

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/272083043>

Finding the beat: A neural perspective across humans and non-human primates

Article in Philosophical Transactions of The Royal Society B Biological Sciences · March 2015
DOI: 10.1098/rstb.2014.0093 · Source: PubMed

CITATIONS
73

READS
270

5 authors, including:



Hugo Merchant
Universidad Nacional Autónoma de México
95 PUBLICATIONS 2,617 CITATIONS

SEE PROFILE



Jessica Adrienne Grah
The University of Western Ontario
94 PUBLICATIONS 3,031 CITATIONS

SEE PROFILE



Laurel Trainor
McMaster University
200 PUBLICATIONS 9,517 CITATIONS

SEE PROFILE



Martin Rohrmeier
École Polytechnique Fédérale de Lausanne
40 PUBLICATIONS 736 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Can monkeys sense the beat or, if not, the regularity of a musical signal? [View project](#)



Development of auditory scene analysis [View project](#)



Review

Cite this article: Merchant H, Grahm J, Trainor L, Rohrmeier M, Fitch WT. 2015 Finding the beat: a neural perspective across humans and non-human primates. *Phil. Trans. R. Soc. B* **370**: 20140093.
<http://dx.doi.org/10.1098/rstb.2014.0093>

One contribution of 12 to a theme Issue
'Biology, cognition and origins of musicality'.

Subject Areas:

neuroscience, computational biology

Keywords:

beat perception, beat synchronization,
neurophysiology, modelling

Author for correspondence:

Hugo Merchant
e-mail: hugomerchant@unam.mx

[†]Present address: Institut für Kunst- und
Musikwissenschaft, Technische Universität
Dresden, August-Bebel-Straße 20, 01219
Dresden, Germany

Finding the beat: a neural perspective
across humans and non-human primates

Hugo Merchant¹, Jessica Grahm², Laurel Trainor³, Martin Rohrmeier^{4,†}
and W. Tecumseh Fitch⁵

¹Instituto de Neurobiología, UNAM, campus Juriquilla, Querétaro 76230, México

²Brain and Mind Institute, and Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5B7

³Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main St. W., Hamilton, Ontario, Canada

⁴Department of Linguistics and Philosophy, MIT Intelligence Initiative, Cambridge, MA 02139, USA

⁵Department of Cognitive Biology, University of Vienna, Althanstrasse 14, Vienna 1090, Austria

Humans possess an ability to perceive and synchronize movements to the beat in music ('beat perception and synchronization'), and recent neuroscientific data have offered new insights into this beat-finding capacity at multiple neural levels. Here, we review and compare behavioural and neural data on temporal and sequential processing during beat perception and entrainment tasks in macaques (including direct neural recording and local field potential (LFP)) and humans (including fMRI, EEG and MEG). These abilities rest upon a distributed set of circuits that include the motor cortico-basal-ganglia–thalamo-cortical (mCBGT) circuit, where the supplementary motor cortex (SMA) and the putamen are critical cortical and subcortical nodes, respectively. In addition, a cortical loop between motor and auditory areas, connected through delta and beta oscillatory activity, is deeply involved in these behaviours, with motor regions providing the predictive timing needed for the perception of, and entrainment to, musical rhythms. The neural discharge rate and the LFP oscillatory activity in the gamma- and beta-bands in the putamen and SMA of monkeys are tuned to the duration of intervals produced during a beat synchronization–continuation task (SCT). Hence, the tempo during beat synchronization is represented by different interval-tuned cells that are activated depending on the produced interval. In addition, cells in these areas are tuned to the serial-order elements of the SCT. Thus, the underpinnings of beat synchronization are intrinsically linked to the dynamics of cell populations tuned for duration and serial order throughout the mCBGT. We suggest that a cross-species comparison of behaviours and the neural circuits supporting them sets the stage for a new generation of neurally grounded computational models for beat perception and synchronization.

1. Introduction

Beat perception is a cognitive ability that allows the detection of a regular pulse (or beat) in music and permits synchronous responding to this pulse during dancing and musical ensemble playing [1,2]. Most people can recognize and reproduce a large number of rhythms and can move in synchrony to the beat by rhythmically timed movements of different body parts (such as finger or foot taps, or body sway). Beat perception and synchronization can be considered fundamental musical traits that, arguably, played a decisive role in the origins of music [1]. A large proportion of human music is organized by a quasi-isochronous pulse and frequently also in a metrical hierarchy, in which the beats of one level are typically spaced at two or three times those of a faster level (i.e. in the most simple Western cases the tempo of one level is 1/2 (march metre) or 1/3 (waltz metre) that of the other), and human listeners can typically synchronize at more than one level of the metrical hierarchy [3,4]. Furthermore, movement on every second versus every third beat of an ambiguous rhythm pattern (one, for example, that can be interpreted as

either a march or a waltz) biases listeners to interpret it as either a march or a waltz, respectively [5]. Therefore, the concept of ‘beat perception and synchronization’ implies both that (i) the beat does not always need to be physically present in order to be ‘perceived’ and (ii) the pulse evokes a particular perceptual pattern in the subject via active cognitive processes. Interestingly, humans do not need special training to perceive and motorically entrain to the beat in musical rhythms; rather it appears to be a robust, ubiquitous and intuitive behaviour. Indeed, even young infants perceive metrical structure [6], and if an infant is bounced on every second beat or on every third beat of an ambiguous rhythm pattern, the infant is biased to interpret the metre of the auditory rhythm in a manner consistent with how they were moved to it. Thus, although rhythmic entrainment is a complex phenomenon that depends on a dynamic interaction between the auditory and motor systems in the brain [7,8], it emerges very early in development without special training [9].

Recent studies support the notion that the timing mechanisms used in the brain depend on whether the time intervals in a sequence can be timed relative to a steady beat (relative, or beat-based, timing) or not (absolute, or duration-based) timing [10–12]. In relative timing, time intervals are measured relative to a regular perceived beat [12], to which individuals are able to entrain. In absolute timing, the absolute duration of individual time intervals is encoded discretely, like a stopwatch, and no entrainment is possible. In this regard, the recent ‘gradual audio-motor hypothesis’ suggests that the complex entrainment abilities of humans seem to have evolved gradually across primates, with a duration-based timing mechanism present across the entire primate order [13,14] and a beat-based mechanism that (i) is most developed in humans, (ii) shows some but not all the properties in monkeys and (iii) is present at an intermediate level chimpanzees [7]. For example, a myriad of studies have demonstrated that humans rhythmically entrain to isochronous stimuli with almost perfect tempo and phase matching [15]. Tempo/period matching means that the period of movement precisely equals the musical beat period. Phase matching means that rhythmic movements occur near or at the onset times of musical beats.

Macaques were able to produce rhythmic movements with proper tempo matching during a synchronization–continuation task (SCT), where they tapped on a push-button to produce six isochronous intervals in a sequence, three guided by stimuli, followed by three internally timed (without the sound) intervals [16]. Macaques reproduced the intervals with only slight underestimations (approx. 50 ms), and their inter-tap interval variability increased as a function of the target interval, as does human subjects’ variability in the same task [16,17]. Crucially, these monkeys produce isochronous rhythmic movements by temporalizing the pause between movements and not the movements’ duration [18], reminiscent of human results [19]. These observations suggest that monkeys use an explicit timing strategy to perform the SCT, where the timing mechanism controls the duration of the movement pauses, which also trigger the execution of stereotyped pushing movements across each produced interval in the rhythmic sequence. On the other hand, however, the taps of macaques typically occur about 250 ms after stimulus onset, whereas humans show asynchronies close to zero (perfect phase matching) or even anticipate stimulus onset, moving slightly ahead of the beat [16]. The positive asynchronies in monkeys are shorter than their reaction times in a control task with random inter-stimulus intervals, suggesting

that monkeys do have some temporal prediction capabilities during SCT, but that these abilities are not as finely developed as in humans [16]. Subsequent studies have shown that monkeys’ asynchronies can be reduced to about 100 ms with a different training strategy [7] and that monkeys show tempo matching to periods between 350 and 1000 ms, similar to what has been seen in humans [7,20]. Hence, it appears that macaques possess some but not all the components of the brain machinery used in humans for beat perception and synchronization [21,22].

In order to understand how motoric entrainment to a musical beat is accomplished in the brain, it is important to compare neurophysiological and behavioural data across humans and non-human primates. Because humans appear to have uniquely good abilities for beat perception and synchronization, it is important to examine the basis of these abilities using functional magnetic resonance imaging (fMRI), electroencephalogram (EEG) and magnetoencephalogram (MEG) data. Using these techniques, a complementary set of studies have described the neural circuits and the potential neural mechanisms engaged during both rhythmic entrainment and beat perception in the humans. In addition, the recording of multiple-site extracellular signals in behaving monkeys has provided important clues about the neural underpinnings of temporal and sequential aspects of rhythmic entrainment. Most of the neural data in monkeys have been collected during the SCT task described above. Hence, spiking responses of cells and local field potential (LFP) recordings of monkeys during the synchronization phase of the SCT must be compared and contrasted to neural studies of beat perception and synchronization in humans, which use more macroscopic techniques such as EEG and brain imaging. In this paper, we review both human and monkey data and attempt to synthesize these different neuronal levels of explanation. We end by providing a set of desiderata for neurally grounded computational models of beat perception and synchronization.

2. Functional imaging of beat perception and entrainment in humans

Although studies with humans have used the SCT task, generally with isochronous sequences, humans can spontaneously entrain to non-isochronous sequences, as well, if they have a temporal structure that induces beat perception [23–25]. Sequences that induce a beat are often termed metric and activity to these sequences can be compared with activity to sequences that do not (non-metric). Different researchers use somewhat different heuristics and terminology for creating metric and non-metric sequences, but the underlying idea is similar: simple metric rhythms induce clear beat perception, complex metric rhythms less so and non-metric rhythms not at all.

During beat perception and synchronization, activity is consistently observed in several brain areas. Subcortical structures include the cerebellum, the basal ganglia (most often the putamen, also caudate nucleus and globus pallidus) and thalamus, and cortical areas include the supplementary motor area (SMA) and pre-SMA, premotor cortex (PMC), as well as auditory cortex [12,26–37]. Less frequently, ventrolateral prefrontal cortex (VLPFC, sometimes labelled as anterior insula) and inferior parietal cortex activations are observed [30,34,38,39]. These areas are depicted in figure 1. Although the specific role of each area is still emerging, evidence is accumulating for distinctions between particular areas and

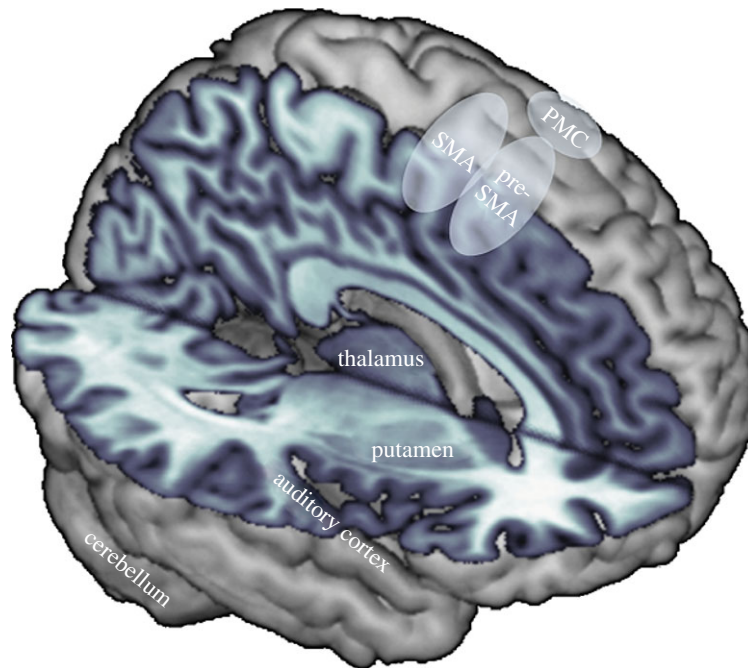


Figure 1. Brain areas commonly activated in functional magnetic resonance studies of rhythm perception (cut-out to show mid-sagittal and horizontal planes, tinted purple). Besides auditory cortex and the thalamus, many of the brain areas of the rhythm network are traditionally thought to be part of the motor system. SMA, supplementary motor area; PMC, premotor cortex. (Online version in colour.)

networks. The basal ganglia and SMA appear to be involved in beat perception, whereas the cerebellum does not. Visual beat perception may be mediated by similar mechanisms to auditory beat perception. Beat perception also leads to greater functional connectivity, or interaction, between auditory and motor areas, particularly for musicians. We now consider this evidence in more detail.

Several fMRI studies indicate that the cerebellum is more important for absolute than relative timing. During various perceptual judgement tasks, the cerebellum is more active for non-metric than metric rhythms [12,32]. Similarly, when participants tap along to sequences, cerebellar activity increases as metric complexity increases and the beat becomes more difficult to perceive [34]. The cerebellum also responds more during learning of non-metric than metric rhythms [40]. The fMRI findings are supported by findings from other methods: deficits in encoding of single durations, but not of metric sequences, occur when cerebellar function is disrupted, either by disease [41] or through transcranial magnetic stimulation [42]. Thus, although the cerebellum is commonly activated during rhythm tasks, the evidence indicates it is involved in absolute, not relative, timing and therefore does not play a significant role in beat perception or entrainment.

By contrast, relative timing appears to rely on the basal ganglia, specifically the putamen, and the SMA, as simple metric rhythms, compared with complex or non-metric rhythms, elicit greater putamen and SMA activity across various perception and production tasks [12,30,33,43,44]. For complex rhythms, SMA and putamen activity can be observed when a beat is eventually induced by several repetitions of the rhythm [45]. Importantly, increases in putamen and SMA activity during simple metric compared with non-metric rhythms cannot be attributed to non-metric rhythms simply being more difficult: even when task difficulty is manipulated to equate performance, greater putamen and SMA activity is still evident for simple metric rhythms [30]. Furthermore, the greater activity occurs even when participants are not instructed to attend to the rhythms [26], or when they attend to non-temporal

aspects of the stimuli such as loudness [43] or pitch [32]. Thus, greater putamen activity cannot be attributed to metric rhythms simply facilitating performance on temporal tasks. Finally, when the basal ganglia are compromised, as in Parkinson's disease, discrimination performance with simple metric rhythms is selectively impaired. Overall, these findings indicate that the basal ganglia not only respond during beat perception, but are crucial for normal beat perception to occur.

Interestingly, basal ganglia activity does not appear to correlate with the speed of the beat that is perceived [28], instead showing maximal activity around 500 to 700 ms [46,47]. Findings from behavioural work suggest that beat perception is maximal at a preferred beat period near 500 ms [48,49]. Therefore, the basal ganglia are not simply responding equally to any perceived temporal regularity in the stimuli, but are most responsive to regularity at the frequency that best induces a sense of the beat.

Most fMRI studies of beat perception use auditory stimuli, as beat perception and synchronization occur more readily with auditory than visual stimuli [23,50–55]. Thus, it is unclear whether visual beat perception would also be mediated by basal ganglia structures. One study found that presenting visual sequences with isochronous timing elicited greater basal ganglia activity than randomly timed stimuli [56], although it is unclear if participants truly perceived a beat in the visual condition. Another way to induce visual beat perception is for visual rhythms to be 'primed', by earlier presentations of the same rhythm in the auditory modality. When metric visual rhythms are perceived after auditory rhythms, putamen activity is greater than when visual rhythms are presented without auditory priming. Moreover, the amount of that increase predicts whether a beat is perceived in the visual rhythm [44]. This suggests that when an internal representation of the beat is induced during the auditory presentation, the beat can be continued in subsequently presented visual rhythms, and this visual beat perception is mediated by the basal ganglia.

In addition to measuring regional activity, fMRI can be used to assess functional connectivity, or interactions, between brain

areas. During beat perception, greater connectivity is observed between the basal ganglia and cortical motor areas, such as the SMA and PMC [32]. Furthermore, connectivity between PMC and auditory cortex was found in one study to increase as the salience of the beat in an isochronous sequence increased [29]. However, a later study using metric, not isochronous, sequences found that auditory–premotor activity increased as metric complexity increased (i.e. connectivity increased as perception of the beat decreased) [34]. Thus, further clarification is needed about the role of auditory–premotor connectivity in isochronous versus metric sequences. Musical training is also associated with greater connectivity between motor and auditory areas. During a synchronization task, musicians showed more bilateral patterns of auditory–premotor connectivity than non-musicians [28]. Musicians also showed greater auditory–premotor connectivity than non-musicians during passive listening, when no movement was required [32]. Thus, beat perception increases functional connectivity both within the motor system and between motor and auditory systems. One hypothesis is that increased auditory–premotor connectivity might be important for integrating auditory perception with a motor response [8,28], and perhaps this occurs even if the response is not executed. Interestingly, this hypothesis has been confirmed using EEG and MEG techniques, as reviewed below, suggesting a fundamental role of the audio–motor system in beat perception and synchronization.

Beat perception unfolds over time: initially, when a rhythm is first heard, the beat must be discovered. After beat-finding occurs, an internal representation of the beat rate can be formed, allowing prediction of future beats as the rhythm continues. Two fMRI studies have attempted to determine whether the role of the basal ganglia is finding the beat, predicting future beats, or both [33,34]. In the first study, participants heard multiple rhythms in a row that either did or did not have a beat. Putamen activity was low during the initial presentation of a beat-based rhythm, during which participants were engaged in finding the beat. Activity was high when beat-based rhythms followed one after the other, during which participants had a strong sense of the beat, suggesting that the putamen is more involved in predicting than finding the beat. The suggestion that the putamen and SMA are involved in maintaining an internal representation of beat intervals is supported by findings of greater putamen and SMA activation during the continuation phase, and not the synchronization phase, during synchronization–continuation tasks [35,57] similar to those described for macaques in the next section [17]. Patients with SMA lesions also show a selective deficit in the continuation phase but not the synchronization phase of the synchronization–continuation task [58]. However, a second fMRI study compared activation during the initial hearing of a rhythm, during which participants were engaged in beat-finding, to subsequent tapping of the beat as they heard the rhythm again. In contrast to the previous study, putamen activity was similar during finding and synchronized tapping [34]. These different fMRI findings may result from different experimental paradigms, stimuli, or analyses, but more research will be needed to determine the role of the basal ganglia in beat-finding and beat prediction.

In summary, a network of motor and auditory areas respond not only during tapping in synchrony with a beat, but also during perception of sequences that have a beat, to which one could synchronize, or entrain. Beat perception elicits greater activity than perception of rhythms without a beat

in a subset of these motor areas, including the putamen and SMA. Motor and auditory areas also exhibit greater coupling during synchronization to and perception of the beat. A viable mechanism for this coupling may be oscillatory responses, not directly measurable with fMRI, but readily observed using EEG or MEG (discussed below).

3. Oscillatory mechanisms underlying rhythmic behaviour in humans: evidence from EEG and MEG

The perception of a regular beat in music can be studied in human adults and newborns [6,59], as well as in non-human primates [60], using a particular event-related brain potential (ERP) called mismatch negativity (MMN). MMN is a preattentive brain response reflecting cortical processing of rare unexpected ('deviant') events in a series of ongoing 'standard' events [61]. Using a complex repeating beat pattern, MMN was observed in human adults in response to changes in the pattern at strong beat locations, suggesting extraction of metrical beat structure. Similar results were found in newborn infants, but unfortunately the stimuli used with infants confounded MMN responses to beat changes with a general response to a change in the number of instruments sounding [6], so future studies are needed with newborns in order to verify these results. Honing *et al.* [60] also recorded ERPs from the scalp of macaque monkeys. This study demonstrated that an MMN-like ERP component could be measured in rhesus monkeys, both for pitch deviants and unexpected omissions from an isochronous tone sequence. However, the study also showed that macaques are not able to detect the beat induced by a varying complex rhythm, while being sensitive to the rhythmic grouping structure. Consequently, these results support the notion of different neural networks being active for interval- and beat-based timing, with a shared interval-based timing mechanism across primates as discussed in the Introduction [7]. A subsequent oddball behavioural study gave additional support for the idea that monkeys can extract temporal information for isochronous beat sequences but not from complex metrical stimuli. This study showed that macaques can detect deviants when they occur in a regular (isochronous) rather than an irregular sequence of tones, by showing changes of gaze and facial expressions to these deviants [62]. However, the authors found that the sensitivity for deviant detection is more developed in humans than monkeys. Humans detect deviants with high accuracy not only for the isochronous tone sequences but also in more complex sequences (with an additional sound with a different timbre), whereas monkeys do not show sensitivity to deviants under these conditions [62]. Overall, these findings suggest that monkeys have some capabilities for beat perception, particularly when the stimuli are isochronous, corroborating the hypothesis that the complex entrainment abilities of humans have evolved gradually across primates, and with a primordial beat-based mechanism already present in macaques [7].

The MMN has also been used to examine another interesting feature of musical rhythm perception, namely that across cultures the basic beat tends to be laid down by bass-range instruments. Hove *et al.* [63] showed that when two tones are played simultaneously in a repeated isochronous rhythm, deviations in the timing of the lower-pitched tone elicits a

larger MMN response compared with deviations in the timing of the higher pitched tone. Furthermore, the results are consistent with behavioural responses. When one tone in an otherwise isochronous sequence occurs too early, and people are tapping along to the beat, they tend to tap too early on tone following the early one, presumably as a form of error correction. Hove *et al.* [63] found that people adjusted the timing of their tapping more when the lower tone was too early compared with when the higher tone was too early. Finally, using a computer model of the peripheral auditory system, these authors showed that this low-voice superiority effect for timing has a peripheral origin in the cochlea of the inner ear. Interestingly, this effect is opposite to that for detecting pitch changes in two simultaneous tones or melodies [64,65], where larger MMN responses are found for pitch deviants in the higher than in the lower tone or melody, an effect found in infants as well as adults [66,67]. The high-voice superiority effect also appears to depend on nonlinear dynamics in the cochlea [68]. It remains unknown as to whether macaques also show a high-voice superiority effect for pitch and a low-voice superiority effect for timing.

In humans, not only do we extract the beat structure from musical rhythms but also doing so propels us to move in time to the extracted beat. In order to move in time to a musical rhythm, the brain must extract the regularity in the incoming temporal information and predict when the next beat will occur. There is increasing evidence across sensory systems, using a range of tasks, that predictive timing involves interactions between sensory and motor systems, and that these interactions are accomplished by oscillatory behaviour of neuronal circuits across a number of frequency bands, particularly the delta (1–3 Hz) and beta (15–30 Hz) bands [69–72]. In general, oscillatory behaviour is thought to reflect communication between different brain regions and, in particular, influences of endogenous ‘top-down’ processes of attention and expectation on perception and response selection [73–75]. Importantly, such oscillatory behaviour can be measured in humans using the EEG and MEG.

The optimal range for the perception of musical tempos is 1–3 Hz [76], which coincides with the frequency range of neural delta oscillations. Indeed, neural oscillations in the delta range have been shown to phase align with the tempo of incoming stimuli for musical rhythms [77,78] as well as for stimuli with less regular rhythms, such as speech [79]. Several researchers have recently suggested that temporal prediction is actually accomplished in the motor system, perhaps through some sort of movement simulation [69,80–83], and that this information feeds back to sensory areas (corollary discharges or efference copies) to enhance processing of incoming information at particular points in time. In particular, temporal expectations appear to align the phase of delta oscillations in sensory cortical areas, such that response times to stimuli that happen to be presented at large delta oscillation peaks are processed more quickly [72] and more accurately [80] than stimuli presented at other times. This phase alignment of delta rhythms appears to provide a neural instantiation of dynamic attending theory proposed by Large & Jones [3], whereby attention is drawn to particular points in time, and stimulus processing at those points is enhanced.

Metrical structure is derived in the brain based not only on periodicities in the input rhythm but also on expectations for regularity. Thus, in isochronous sequences of identical tones,

some tones (e.g. every second or every third tones) can be perceived as accented. And beat locations can be perceived as metrically strong even in the presence of syncopation, in which the most prominent physical stimuli occur off the beat. Interestingly, phase alignment of delta oscillations in auditory cortex reflects the metrical interpretation of input rhythms rather than simply the periodicities in the stimulus [77,78] as predicted by resonance theory [84]. For example, when listeners are presented with a 2.4 Hz sequence of tones, a component at 2.4 Hz can be seen in the measured EEG response [77]. But when listeners imagine strong beats every second tone, oscillatory activity in the EEG at 1.2 Hz is also evident, whereas when they imagine strong beats every third tone, EEG oscillations at 0.8 and 1.6 Hz (harmonic of the ternary metre) are evident.

Neural oscillations in the beta range have also been shown to play an important role in predictive timing [70,71,85]. For example, Iversen *et al.* [85] showed that beta but not gamma (30–50 Hz) oscillations evoked by tones in a repeating pattern were affected by whether or not listeners imagined them as being on strong or weak beats. Interestingly, beta oscillations have long been associated with motor processes [86]. For example, beta amplitude decreases during motor planning and movement, recovering to baseline once the movement is completed [87–91].

Of importance in the present context are the findings that beta oscillations are crucial for predictive timing in auditory beat processing and that beta oscillations involve interactions between auditory and motor regions [70,71]. Fujioka *et al.* [71] recorded the MEG while participants listened without attending (while watching a silent movie) to isochronous beat sequences at three different tempos in the delta range (2.5, 1.7 and 1.3 Hz). Examining responses from auditory cortex, they found that the power of induced (non-phase-locked) activity in the beta-band decreased after the onset of each beat and reached a minimum at approximately 200 ms after beat onset regardless of tempo. However, the timing of the beta power rebound depended on the tempo, such that maximum beta-band power was reached just prior to the onset of the next beat (figure 2). Thus, the beta rebound tracked the tempo of the beat and predicted the onset of the following beat. While little developmental work has yet been done, one recent study indicates that similar beta-band fluctuations can be seen in children, at least for slower tempos [92]. Another study [70] found that if a beat in a sequence is omitted, the decrease in beta power does not occur, suggesting that the decrease in beta power might be tied to the sensory input, whereas the rebound in beta power is predictive of the expected onset of the next beat and is internally generated. The results suggest, further, that oscillations in beta and delta frequencies are connected, with beta power fluctuating according to the phase of delta oscillations, consistent with previous work [93].

To examine the role of the motor cortex in beta power fluctuations, Fujioka *et al.* [71] looked across the brain for regions that showed similar fluctuations in beta power as in the auditory cortices and found that this pattern of beta power fluctuation occurred across a wide range of motor areas. Interestingly, the beta power fluctuations appear to be in opposite phase in auditory and motor areas, which is suggestive that the activity might reflect some kind of sensorimotor loop connecting auditory and motor regions. This is consistent with a recent study in the mouse showing that axons from M2 synapse with neurons in deep and superficial layers of auditory cortex,

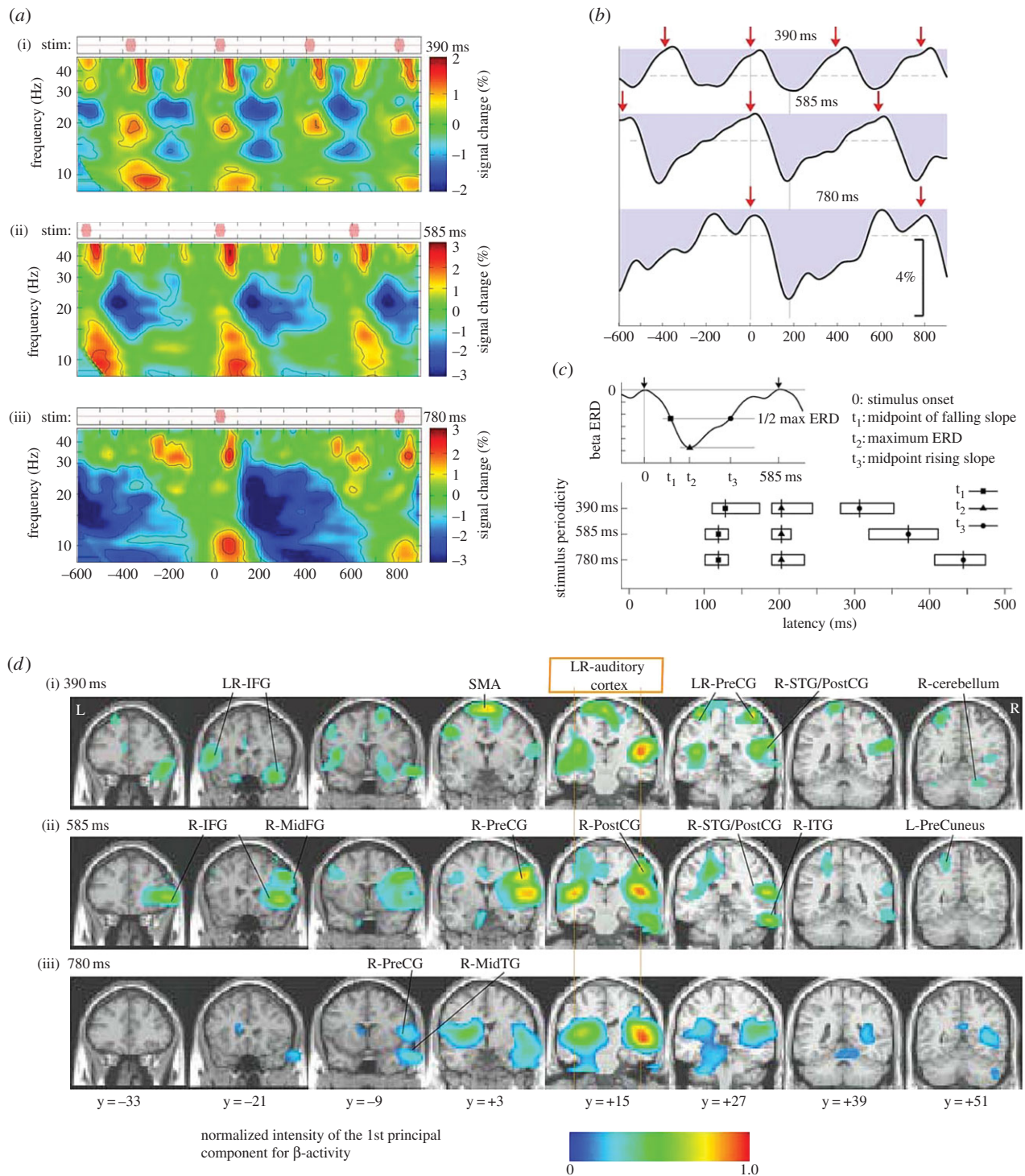


Figure 2. Induced neuromagnetic responses to isochronous beat sequences at three different tempos. (a) Time-frequency plots of induced oscillatory activity in right auditory cortex between 10 and 40 Hz in response to a fast (390 ms onset-to-onset; upper plot), moderate (585 ms; middle plot) and slow (780 ms; lower plot) tempo ($n = 12$). (b) The time courses of oscillatory activity in the beta-band (20–22 Hz) for the three tempos, showing beta desynchronization immediately after stimulus onset (shown by red arrows), followed by a rebound with timing predictive of the onset of the next beat. The dashed horizontal lines indicate the 99% confidence limits for the group mean. (c) The time at which the beta desynchronization reaches half power (squares) and minimum power (triangles), and the time at which the rebound (resynchronization) reaches half power (circles). The timing of the desynchronization is similar across the three stimulus tempos, but the time of the rebound resynchronization depends on the stimulus tempo in a predictive manner. (d) Areas across the brain in which beta-activity is modulated by the auditory stimulus, showing involvement of both auditory cortices and a number of motor regions. (Adapted from Fujioka *et al.* [71].)

and have a largely inhibitory effect on auditory cortex [94]. This is interesting in that interactions between excitatory and inhibitory neurons often give rise to neural oscillations [95]. Furthermore, axons from these same M2 neurons also extend to several subcortical areas important for auditory processing. Finally, Nelson *et al.* [94] found that stimulation of M2 neurons affected the activity of neurons in auditory cortex. Thus, the

connections described by Nelson *et al.* [94] probably reflect the circuits giving rise to the delta-frequency oscillations in beta-band power described by Fujioka *et al.* [71]. Because the task of Fujioka [71] involved listening without attention or motor movement, the results indicate that motor involvement in beat processing is obligatory and may provide a rationale for why music makes people want to move to the beat.

In sum, research is converging that auditory and motor regions connect through oscillatory activity, particularly at delta and beta frequencies, with motor regions providing the predictive timing needed for the perception of, and entrainment to, musical rhythms.

4. Neurophysiology of rhythmic behaviour in monkeys

The study of the neural basis of beat perception in monkeys and its comparison with humans started just recently, such that the first attempts are the previously described MMN experiments in macaques. As far as we know, no neurophysiological study has been performed on a beat perception task in behaving monkeys. By contrast, a complete set of single cell and LFP experiments have been carried out in monkeys during the execution of the SCT. A critical aspect of these studies is that they have been performed in cortical and subcortical areas of the circuit for beat perception and synchronization described in fMRI and MEG studies in humans, specifically in the SMA/pre-SMA, as well as in the putamen of behaving monkeys. These areas are deeply interconnected and are fundamental processing nodes of the motor cortico-basal-ganglia–thalamo-cortical (mCBGT) circuit across all primates [96]. Hence, direct inferences of the functional associations between the different levels of organization measured in this circuit with diverse techniques can be performed across species. This is particularly true for data collected during the synchronization phase of the SCT in monkeys and for human studies of beat synchronization. In addition, cautious generalizations can also be made to the beat perception experiments in humans, as its mechanisms seem to have a large overlap with beat synchronization, as described in the previous two sections.

The signals recorded from extracellular recordings in behaving animals can be filtered to obtain either single cell action potentials or LFPs. Action potentials last approximately 1 ms and are emitted by cells in spike trains, whereas LFPs are complex signals determined by the input activity of an area in terms of population excitatory and inhibitory postsynaptic potentials [97]. Hence, LFPs and spike trains can be considered as the input and output stages of information processing, respectively. Furthermore, LFPs are a local version of the EEG that is not distorted by the meninges and the scalp and that is a signal that provides important information about the input–output processing inside local networks [97]. Now, the analysis of putamenal LFPs in monkeys performing the SCT revealed an orderly change in the power of transient modulations in the gamma (30–70 Hz) and beta (15–30 Hz) as a function of the duration of the intervals produced in the SCT (figure 3) [98]. The burst of LFP oscillations showed different preferred intervals so that a range of recording locations represented all the tested durations. These results suggest that the putamen contains a representation of interval duration, where different local cell populations oscillate in the beta- or gamma-bands for specific intervals during the SCT. Therefore, the transient modulations in the oscillatory activity of different cell ensembles in the putamen as a function of tempo can be part of the neural underpinnings for beat synchronization. Indeed, LFPs tuned to the interval duration in the synchronization phase of the SCT could be considered an empirical substrate of the neural resonance hypothesis that suggests

that perception of pulse and metre result from rhythmic bursts of high-frequency neural activity in response to music [76]. Accordingly, high-frequency bursts in the gamma- and beta-bands may enable communication between neural areas in humans, such as auditory and motor cortices, during rhythm perception and production, as mentioned previously [99,100].

In a series of elegant studies, Schroeder and colleagues have shown that when attention is allocated to auditory or visual events in a rhythmic sequence, delta oscillations of primary visual and auditory cortices of monkeys are entrained (i.e. phase-locked) to the attended modality [72,101]. As a result of this entrainment, the neural excitability and spiking responses of the circuit have the tendency to coincide with the attended sensory events [72]. Furthermore, the magnitude of the spiking responses and the reaction time of monkeys responding to the attended modality in the rhythmic patterns of visual and auditory events are correlated with the delta phase entrainment in a trial by trial basis [72,101]. Finally, attentional modulations for one of the two modalities is accompanied by large-scale neuronal excitability shifts in the delta band across a large circuit of cortical areas in the human brain, including primary sensory and multisensory areas, as well as premotor and prefrontal areas [74]. Therefore, these authors raised the hypothesis that the attention modulation of delta phase-locking across a large cortical circuit is a fundamental mechanism for sensory selection [102]. Needless to say, the delta entrainment across this large cortical circuit should be present during the SCT; however, the putamenal LFPs in monkeys did not show modulations in the delta band during this task. It is known that delta oscillations have their origin in thalamo-cortical interactions [103] and that play a critical role in the long-range interactions between cortical areas in behaving monkeys [104]. Hence, a possible explanation for this apparent discrepancy is that delta oscillations are part of the mechanism of information flow within cortical areas but not across the mCBGT circuit, which in turn could use interval tuning in the beta- and gamma-bands to represent the beat. Subcortical recordings in the relay nuclei of the mCBGT circuit in humans performing beat perception and synchronization could test this hypothesis.

Surprisingly, the transient changes in oscillatory activity also showed an orderly change as a function of the phase of the task, with a strong bias towards the synchronization phase for the gamma-band (when aligned to the stimuli), whereas there is a strong bias towards the continuation phase for the beta-band (when aligned to the tapping movements) [98]. These results are consistent with the notion that gamma-band oscillations predominate during sensory processing [105] or when changes in the sensory input or cognitive set are expected [86]. Thus, the local processing of visual or auditory cues in the gamma-band during the SCT may serve for binding neural ensembles that processes sensory and motor information within the putamen during beat synchronization [98]. However, these findings seem to be at odds with the role of the beta oscillations in connecting the predictive timing signals from human motor regions to auditory areas needed for the perception of, and entrainment to, musical rhythms. In this case, it is also critical to consider that gamma oscillations in the putamen during beat synchronization seem to be local and more related to sensory cue selection for the execution of movement sequences [106,107], rather than associated to the establishment of top-down predictive signal between motor and auditory cortical areas.

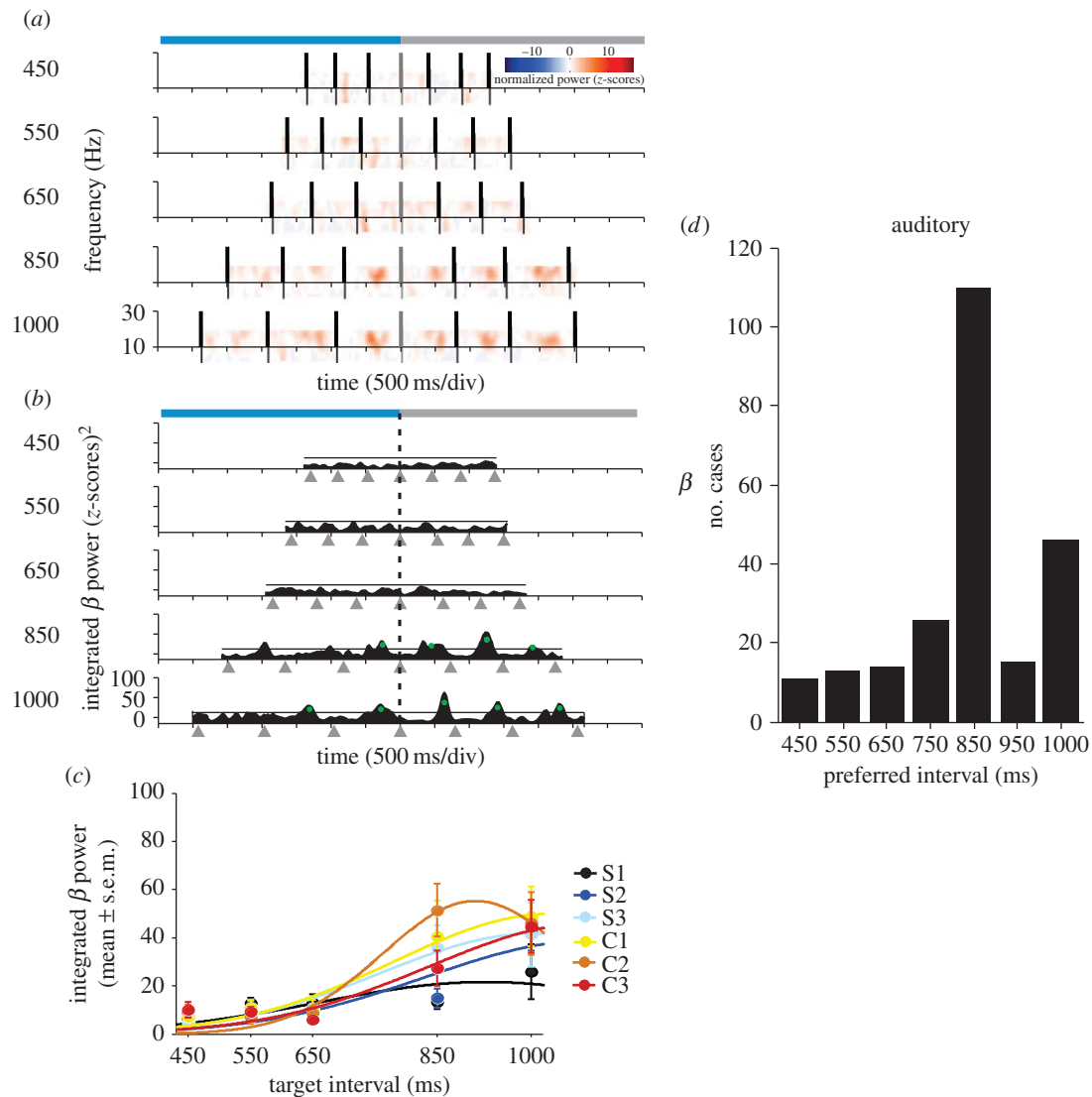


Figure 3. Time-varying modulations in the beta power of an LFP signal that show selectivity to interval duration. (a) Normalized (z-scores, colour-code) spectrograms in the beta-band. Each plot corresponds to the target interval indicated on the left. The horizontal axis is the time during the performance of the task. The times at which the monkey tapped on the push-button are indicated by black vertical bars, and all the spectrograms are aligned to the last tap of the synchronization phase (grey vertical bar). Light-blue and grey horizontal bars at the top of the panel represent the synchronization and continuation phases, respectively. (b) Plots of the integrated power time-series for each target duration. Tap times are indicated by grey triangles below the time axis. Green dots correspond to power modulations above the 1 s.d. threshold (black solid line) for a minimum of 50 ms across trials. The vertical dotted line indicates the last tap of the synchronization phase. (c) Interval tuning in the integrated beta power. Dots are the mean \pm s.e.m. and lines correspond to the fitted Gaussian functions. Tuning functions were calculated for each of the six elements of the task sequence (S1–S3 for synchronization and C1–C3 for continuation) and are colour-coded (see inset). (d) Distribution of preferred interval durations for all the recorded LFPs in the beta-band with significant tuning for the auditory condition. (Adapted from Bartolo *et al.* [98].)

The beta-band activity in the putamen of non-human primates showed a preference for the continuation phase of the SCT, which is characterized by the internal generation of isochronic movements. Consequently, the global beta oscillations could be associated with the maintenance of a rhythmic set and the dominance of an endogenous reverberation in the mCBGT circuit, which in turn could generate internally timed movements and override the effects of external inputs (figure 4) [98]. Comparing the experiments in humans and macaques, it is clear that in both species the beta-band is deeply involved in prediction and the internal set linked with the processing of regular events (isochronic or with a particular metric) across large portions of the brain.

The functional properties of cell spiking responses in the putamen and SMA were also characterized in monkeys performing the SCT. Neurons in these areas show a graded modulation in discharge rate as a function of interval

duration in the SCT [98,108,109,111,112]. Psychophysical studies on learning and generalization of time intervals predicted the existence of cells tuned to specific interval durations [113,114]. Indeed, large populations of cells in both areas of the mCBGT are tuned to different interval durations during the SCT, with a distribution of preferred intervals that covers all durations in the hundreds of milliseconds, although there was a bias towards long preferred intervals [109] (figure 3d). The bias towards the 800 ms interval in the neural population could be associated with the preferred tempo in rhesus monkeys, similar to what has been reported in humans for the putamen activation at the preferred human tempo of 500 ms [46,47]. Thus, the interval tuning observed in the gamma and beta oscillatory activity of putaminal LFPs is also present in the discharge rate of neurons in the SMA and the putamen, and quite probably across all the mCBGT circuit. The tuning association between

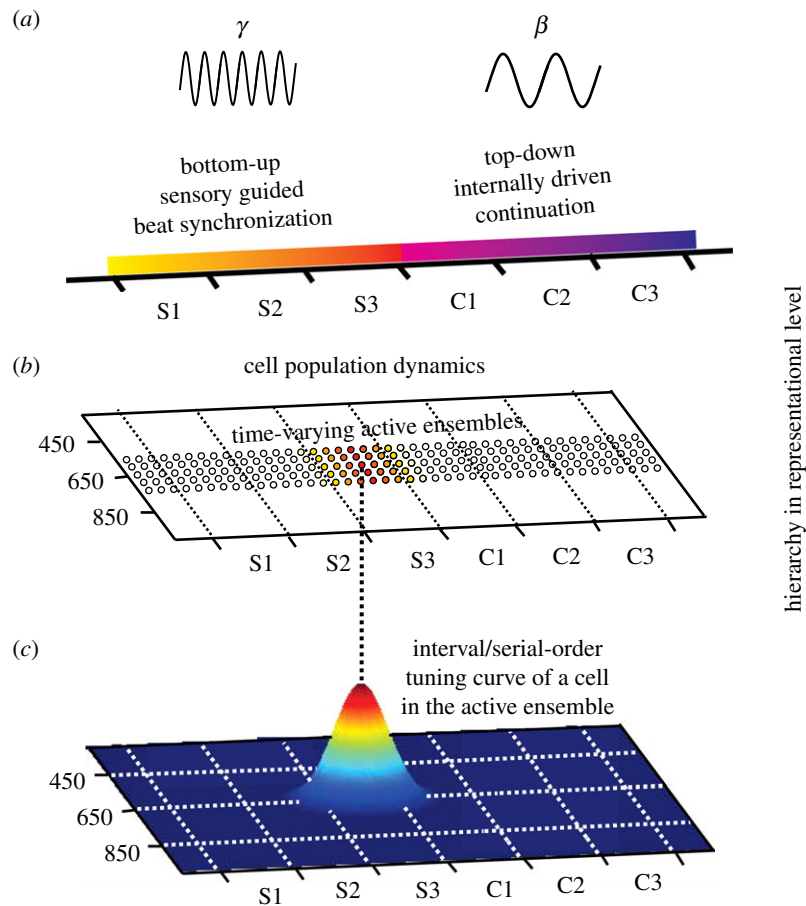


Figure 4. Multiple layers of neuronal representation for duration, serial order and context in which the synchronization–continuation task is performed. (a) LFP activity that switches from gamma-activity during sensory driven (bottom-up) beat synchronization, to beta-activity during the internally generated (top-down) continuation phase. (b) Dynamic representation of the sequential and temporal structure of the SCT. Small ensembles of interconnected cells are activated in a consecutive chain of neural events. (c) A duration/serial-order tuning curve for a cell that is part of the dynamically activated ensemble on top. Dotted horizontal line links the tuning curve of the cell with the neural chain of events in (b). (Adapted from Merchant *et al.* [92,108–110].)

spiking activity and LFPs is not always obvious across the central nervous system, because the latter is a complex signal that depends on the following factors: the input activity of an area in terms of population excitatory and inhibitory postsynaptic potentials; the regional processing of the microcircuit surrounding the recording electrode; the cytoarchitecture of the studied area; and the temporally synchronous fluctuations of the membrane potential in large neuronal aggregates [97]. Consequently, the fact that interval tuning is ubiquitous across areas and neural signals during the SCT underlines the role of this population signal in encoding tempo during beat synchronization and probably during beat perception too.

Beat perception and synchronization not only have a predictive timing component, but also are immersed in a sequence of sensory and motor events. In fact, the regular and repeated cycles of sound and movement during beat synchronization can be fully described by their sequential and temporal information. The neurophysiological recordings in SMA underscored the importance of sequential encoding during the SCT. A large population of cells in this area showed response selectivity to the sequential organization of the SCT, namely, they were tuned to one of the six serial order of the SCT (three in the synchronization and three in the continuation phase; figure 3c) [109]. Again, all the possible preferred serial orders are represented across cell populations. Cell tuning is an encoding mechanism used by the cerebral cortex to represent different sensory, motor and cognitive

features [115], which include the duration of the intervals and the serial order of movements produced rhythmically. These signals must be integrated as a population code, where the cells can vote in favour of their preferred interval/serial order to generate a neural ‘tag’. Hence, the temporal and sequential information is multiplexed in a cell population signal across the mCBGT that works as the notes of a musical score in order to define the duration of the produced interval and its position in the learned SCT sequence [111].

Interestingly, the multiplexed signal for duration and serial order is quite dynamic. Using encoding and decoding algorithms in a time-varying fashion, it was found that SMA cell populations represent the temporal and sequential structure of periodic movements by activating small ensembles of interconnected neurons that encode information in rapid succession, so that the pattern of active tuned neurons changes dramatically within each interval (figure 3b, top) [110]. The progressive activation of different ensembles generates a neural ‘wave’ that represents the complete sequence and duration of produced intervals during an isochronic tapping task such as the SCT [110]. This potential anatomofunctional arrangement should include the dynamic flow of information inside the SMA and through the loops of the mCBGT. Thus, each neuronal pool represents a specific value of both features and is tightly connected, through feed-forward synaptic connections, to the next pool of neurons, as in the case of synfire chains described in neural-network simulations [116].

A key feature of the SMA neural encoding and decoding for duration and serial order is that its width increased as a function of the target interval during the SCT [110]. The time window in which tuned cells represent serial order, for example, is narrow for short interval durations and becomes wider as the interval increases [110]. Thus, duration and serial order are represented in relative terms in SMA, which as far as we know, is the first single cell neural correlate of beat-based timing and is congruent with the functional imaging studies in the mCBGT circuit described above [12].

The investigation of the neural correlates of beat perception and synchronization in humans has stressed the importance of the top-down predictive connection between motor areas and the auditory cortex. A natural following step in the monkey neurophysiological studies should be the simultaneous recording of premotor and auditory areas during tapping synchronization and beat perception tasks. Initial efforts have been started in this direction. However, it is important to emphasize that monkeys show a large bias towards visual rather than auditory cues to drive their tapping behaviour [7,16,117], which contrast with the strong bias towards the auditory modality in humans during music and dance [1,7]. In fact, during the SCT, humans show smaller temporal variability and better accuracy with auditory rather than visual interval markers [54–56]. It has been suggested that the human perceptual system abstracts the rhythmic-temporal structure of visual stimuli into an auditory representation that is automatic and obligatory [118,119]. Thus, it is quite possible that the human auditory system has a privileged access to the temporal and sequential mechanisms working inside the mCBGT circuit in order to determine the exquisite rhythmic abilities of the *Homo sapiens* [7,21]. The monkey brain seems to not have this strong audio-motor network, as revealed in comparative diffusion tensor imaging (DTI) experiments [96,120]. Hence, it is likely that the audio-motor circuit could be less important than the visuo-motor network in monkeys during beat perception and synchronization. The tuning properties of monkey SMA cells multiplexing the temporal and sequential structure of the SCT were very similar across visual and auditory metronomes [109]. However, unpublished observations have shown that a larger number of SMA cells respond specifically to visual rather than to auditory cues during the SCT [121], giving the first support to the monkeys' stronger association in the visuo-motor system during rhythmic behaviours. A final point regarding the modality issue is that human subjects improve their beat synchronization when static visual cues are replaced by moving visual stimuli [52,122], stressing the need of using visual moving stimuli to cue beat perception and synchronization in non-human primates [22,96].

Taken together, the discussed evidence supports the notion that the underpinnings of beat synchronization are intrinsically linked to the dynamics of cell populations tuned for duration and serial order throughout the mCBGT, which represent information in relative- rather than absolute-timing terms. Furthermore, the oscillatory activity of the putamen measured with LFPs showed that gamma-activity reflects local computations associated with sensory-motor (bottom-up) processing during beat synchronization, whereas beta-activity involves the entrainment of large putaminal circuits, probably in conjunction with other elements of the mCBGT, during internally driven (top-down) rhythmic tapping. A critical question regarding these empirical observations is how the different levels of neural representation interact and are coordinated during beat perception

and synchronization. One approach to answering this question would involve integrating the data reviewed above with the results obtained from neurally grounded computational models.

5. Implications for computational models of beat induction

In the past decades, a considerable variety of models have been developed that are relevant in some way to the core neural and comparative issues in rhythm perception discussed above. Broadly speaking, they differ with respect to their purpose and the level of cognitive modelling (in David Marr's sense of algorithmic, computational and implementational levels [123]). These models include rule-based cognitive models, musical information retrieval (MIR) systems designed to process and categorize music files, and dynamical-systems models. Although we make no attempt to review all such models here in detail, we will make a few general observations about the usefulness of each model type for our central problem in this review: understanding the neural basis and comparative distribution of rhythmic abilities.

Rule-based models (e.g. [124,125]) are intended to provide a very high-level computational description of what rhythmic cognition entails, in terms of both pulse and metre induction (reviewed in [4]). Such models make little attempt to specify how, computationally or neurally, this is accomplished, but they do provide a clear description of what any neurally grounded model of human beat perception and synchronization and metre perception should be able to explain. Thus, Lerdahl & Jackendoff [124] emphasize the need to attribute both pulse and metre to a musical surface and propose various abstract principles in terms of well-formedness and preference rules to accomplish these goals. Longuet-Higgins & Lee [125] emphasize that any model of musical cognition must be able to cope with syncopated rhythms, in which some musical events do not coincide with the pulse or a series of events appear to conflict with the established pulse.

A central point made by many rule-based modellers concerns the importance of both bottom-up and top-down processes in determining the rhythmic inferences made by a listener [126]. As a simple example, the vast majority of drum patterns in popular music have bass drum hits on the down-beat (e.g. beats 1 and 3 of a 4/4 pattern) and snare hits on the upbeats (beats 2 and 4). While not inviolable, this statistical generalization leads listeners familiar with these genres to have strong expectations about how to assign musical events to a particular metrical position, thus using learned top-down cognitive processes to reduce the inherent ambiguity of the musical surface. Similarly, many genres have particular rhythmic tropes (e.g. the clave pattern typical of salsa music) that, once recognized, provide an immediate orientation to the pulse and metre in this style. This presumably allows experienced listeners to infer pulse and metre more rapidly (e.g. after a single measure) than would be possible for a listener relying solely on bottom-up processing. It would be desirable for models to allow this type of top-down processing to occur and to provide mechanisms whereby learning about a style can influence more stimulus-driven bottom-up processing.

A second class of potential models for rhythmic processing comes from the practical world of MIR systems. Despite their commercial orientation, such models must efficiently solve problems similar to those addressed in cognitively orientated

models. Particularly relevant are beat-finding and tempo-estimation algorithms that aim to process recordings and accurately pinpoint the rate and phase of pulses (e.g. [127,128]). Because such models are designed to deal with real recorded music files, successful MIR systems have higher ecological validity than rule-based systems designed to deal only with notated scores or MIDI files. Furthermore, because their goal is practical and performance driven, and backed by considerable research money, engineering expertise and competitive evaluation (such as MIREX), we can expect this range of models to provide a 'menu' of systems able to do in practice what any normal human listener easily accomplishes. While we cannot necessarily expect such systems to show any principled connection to the computations humans in fact use to find and track a musical beat, such models should eventually provide insight into which computational approaches and algorithms do and don't succeed, and about which challenges are successfully addressed by which techniques. Some of these findings should be relevant to cognitive and neural researchers. Unfortunately, and perhaps surprisingly, no MIR system currently available can fully reproduce the rhythmic processing capabilities of a normal human listener [128]. At present, this class of models supports the contention that, despite its intuitive simplicity, human rhythmic processing is by no means trivial computationally.

The third class of models—dynamical-systems-based approaches—is most relevant to our purposes in this paper. Such models attempt to be mathematically explicit and to engage with real musical excerpts (not just scores or other symbolic abstractions). As for the other two classes, there is considerable variety within this class, and we will not attempt a comprehensive review. Rather, we will focus on a class of models developed by Edward Large and his colleagues over several decades, which represent the current state of the art for this model category.

Large and his colleagues [3,76,129,130] have introduced and developed a class of mathematical models of beat induction that can reproduce many of the core behavioural characteristics of human beat and metre perception. Because the model of Large [76] has served as the basis for subsequent improvements and is comprehensively and accessibly described, we describe it briefly here. Large's model is based on a set of nonlinear Hopf oscillators. Such oscillators have two main states—damped oscillation and self-sustained oscillation—where an energy parameter α determines which of these states the system is in. Sustained oscillation is the state of interest because only in this state can an oscillator maintain an implicit beat in the face of temporary silence or conflicting information (e.g. syncopation). In Large's model [76], a set of such nonlinear Hopf oscillators is assembled into a network in which each oscillator has inhibitory connections with the others. This leads to a network where each oscillator competes with the others for activation, and the winner(s) are determined by the rhythmic input signal fed into the entire network. Practically speaking, the Large model [76] had 96 such oscillators with fixed preferred periods spaced logarithmically from 100 ms (600 BPM) to 1500 ms (40 BPM) and thus covering the effective span of human tempo perception. Because each oscillator has its own preferred oscillation rate and a coupling strength to the input, as well as a specific inhibitory connection to the other $N - 1$ oscillators, this model overall has more than N^2 (in this case, more than 9000) free parameters. However, these parameters are mostly set based on *a priori* considerations

of human rhythmic perception (e.g. for preferred pulse rate around 600 ms or the tempo range described above) to avoid 'tweaking' parameters to fit a given dataset. In Large [76], this model was tested using ragtime piano excerpts, and the results compared to experimental data produced by humans tapping to these same excerpts [131]. A good fit was found between human performance and the model's behaviour.

Nonlinear oscillator networks of this sort exhibit a number of desirable properties important for any algorithm-level model of human beat and metre perception. First and foremost, the notion of self-sustained oscillation allows the network, once stimulated with rhythmic input, to persist in its oscillation in the face of noise, syncopation or silence. This is a crucial feature for any cognitive model of beat perception that is both robust and predictive. Second, the networks explored by Large and colleagues can exhibit hysteresis (meaning that the current behaviour of the system depends not only on the current input but on the system's past behaviour). Thus, once a given oscillator or group of oscillators is active, they tend to stay that way unless outcompeted by other oscillators, again providing resistance to premature resetting of tempo or phase inference due on syncopation or short-term deviations. Beyond these beat-based factors, these models also provide for metre perception, in that multiple related oscillators can be, and typically are, active simultaneously. Thus, the system as a whole represents not just the tactus (typically tapping rate) but also harmonics or subharmonics of that rate. Finally, these models have been tested and vetted against multiple sources of data, including both behavioural and more recently neural studies, and generally perform quite well at matching the specific patterns in data derived from humans [130].

Despite these many virtues, these models also have certain drawbacks for those seeking to understand the neural basis and comparative distribution of rhythmic capabilities. Most fundamentally, these models are implemented at a rather generic mathematical level. Because they have been designed to mirror cognitive and behavioural data, neither the oscillator nor the network properties have any direct connection to properties of specific brain regions, neural circuits or measurable neuronal properties. For example, both the human preference for a tactus period around 600 ms, and our species' preference for small integer ratios in metre, are hand-coded into the model, rather than deriving from any more fundamental properties of the brain systems involved in rhythm perception (e.g. refractory periods of neurons or oscillatory time constants of cortical or basal ganglia circuits). Large [76, p. 533] explicitly considers this generic mathematical framework, implemented 'without worrying too much about the details of the neural system', to be a desirable property. However, it makes the model and its parameters difficult to integrate with the converging body of neuroscientific evidence reviewed above.

Thus, for example, we would like to know *why* humans (and perhaps some other animals) prefer to have a tactus around 600 ms. Does this follow from some external property of the body (e.g. resonance frequency of the motor effectors) or some internal neural factor (e.g. preferred frequency for cortico-cortical sensorimotor oscillations, or preferred firing rates of basal ganglia neurons)? If the latter, are the observed behavioural preferences a result of properties of the oscillators themselves (as implied in [130]), of the network connectivity [76], or both? Similarly, a neurally grounded model should eventually be able to derive the human preference for small

integer ratios in metre perception from some more fundamental neural and/or cognitive constraints (e.g. properties of coupling between cortical and basal ganglia circuits) rather than hard-coding them into the model. Ultimately, human cognitive neuroscientists need models that make more specific testable predictions about the neural circuitry, predictions that can be evaluated using modern brain imaging data.

Finally, the apparently patchy comparative distribution of rhythmic abilities among different animal species remains mysterious from the viewpoint of a generic model that, in principle, applies to the nervous system of any species possessing a set of self-sustaining nonlinear oscillators (e.g. virtually any vertebrate species). Why is it that macaques, despite their many similarities with humans, apparently find it so difficult to predictively time their taps to a beat, and show no evidence for grouping events into metrical structures [7,60]? Why is it that the one chimpanzee (out of three tested) who shows any evidence at all for beat perception and synchronization only does so for a single preferred tempo, and does not generalize this behaviour to other tempos [132]? What is it about the brain of parrots or sea lions that, in contrast to these primates, enables them to quite flexibly entrain to rhythmic musical stimuli [133–136]? Why is it that some very small and apparently simple brains (e.g. in fireflies or crickets) can accomplish flexible entrainment, while those of dogs apparently cannot [137]? To address any of these questions requires models more closely tied to measurable properties of animal brains than those currently available.

Another desideratum for the next generation of models of rhythmic cognition would include more explicit treatment of the motor output component observed in entrainment behaviour, to help understand when and why such movements occur. Humans can tap their fingers, nod their heads, tap their feet, sway their bodies or combine such movements during dancing, and all these output behaviours appear to be to some extent cognitively interchangeable. Musicians can do the same with either their voices or their hands and feet. Does this remarkable flexibility mean that rhythmic cognition is primarily a sensory and central phenomenon, with no essential connection to the details of the motor output? With different species, does it matter that in some species (e.g. non-human primates) we study finger taps as the motor output, whereas in others (e.g. birds or pinnipeds) we use head bobs or beak taps? Could other species do better with a vocal response? To what extent do different models predict that output mode should play an important role in determining success or failure in beat perception and synchronization tasks, or in preferred entrainment tempos? While incorporating a motor component in perceptual models does not pose a major computational challenge, it would allow us to evaluate such issues, and perhaps help understand the different roles of cortical, cerebellar and basal ganglia circuits in human and animal rhythmic cognition.

Several features outlined in the review above could provide fertile inspiration for implementational models at the neural level. First, the several independent loops involved in rhythmic cognition have different fundamental properties. Cortico-cortical loops (e.g. connecting premotor and auditory regions) clearly play an important role in rhythm perception, and a key characteristic of long-range connections in cortex is that they are excitatory—pyramidal cells exciting other pyramidal cells in a feedback loop that in principle can lead to uncontrolled positive feedback and seizure. This is avoided in cortex via local inhibitory neurons. Such inhibitory interneurons in turn provide a target for excitatory feedback projections to ‘sculpt’ ongoing

local activity by stimulating local inhibition. This can be conceptualized as high-level motor systems making predictions that bias information processing in ‘lower’ auditory regions, via local inhibitory interneurons that tune local circuit oscillations. In sharp contrast, basal ganglia loops are characterized by pervasive inhibitory connections, where inhibition of inhibition is the rule. The computational character of these two loops is thus fundamentally different, in ways that are likely, if properly modelled, to have predictable consequences in terms of both temporal integration windows and the way in which cortico-cortical and basal ganglia loops interact with and influence one another. The basal ganglia are also a main locus for dopaminergic reward circuitry projections from the midbrain, which inject learning signals into the forebrain loops. Thus, this circuit may be a preferred locus for the rewarding effects of rhythmic entrainment and/or learning of rhythmic patterns.

Clearly, each of these classes of models addresses interesting questions, and one cannot expect any single model to span all of these levels. Models at the rule-based (computational) and algorithmic levels are currently the most mature and have already provided a good scaffolding for empirical research in cognition and behaviour. However, our progress in understanding the neurocomputational basis of rhythm perception in humans and other animals will require a more concerted focus on the computational properties of actual neural circuits, at an implementational level. We hope that the brief review and critique above helps to encourage more attention to modelling at Marr’s implementational level.

6. Conclusion

This paper describes and compares current knowledge concerning the anatomical and functional basis of beat perception and synchronization in human and non-human primates, ending by delineating how this knowledge could be used to build neurally grounded models. Hence, our review provides an integrated panorama across fields that have only been treated separately before [7,23,138,139]. It is clear that the human mCBGT circuit is engaged not only during motoric entrainment to a musical beat but also during the perception of simple metric rhythms. This indicates that the motor system is involved in the representation of the metrical structure of auditory stimuli. Furthermore, this motoric representation is predictive and can induce in auditory cortex an expectation process for metrical stimuli. The predictive signals are conveyed to the sensory areas via oscillatory activity, particularly at delta and beta frequencies. Non-invasive data from humans are complemented by direct recordings of single cell and microcircuit activity in behaving macaques, showing that SMA, the putamen and probably all of the relay nuclei of the mCBGT circuit use different encoding strategies to represent the temporal and sequential structure of beat synchronization. Indeed, interval tuning could be a mechanism used by the mCBGT to represent the beat tempo during synchronization. As for humans, oscillatory activity in the beta-band is deeply involved in generating the internal set used to process regular events. There is a strong consensus that the motor system makes use of multiple levels of neural representation during beat perception and synchronization. However, implementation-level models, more tightly tied to properties of cells and neural circuits, are urgently needed to help describe and make sense of this (still incomplete) empirical information. Dynamical

system approaches to model the neural representations of beat are successful at an algorithmic level, but incorporating single cell intrinsic properties, cell tuning and ramping activity, microcircuit organization and connectivity, and the dynamic communication between cortical and subcortical areas in realistic models would be welcome. The predictions generated by neurally grounded models would help drive further empirical research to bridge the gap across different levels of brain organization during beat perception and synchronization.

Author contributions. H.M. wrote the introduction, the section of neurophysiology of rhythmic behaviour in monkeys, and the conclusions; J.G. wrote the section on the functional imaging of beat perception and entrainment in humans; L.T. wrote the section on the oscillatory

mechanisms underlying rhythmic behaviour in humans; M.R. and T.F. wrote the section on the computational models for beat perception and synchronization, and T.F. edited the whole manuscript. All authors contributed to the ideas and framework contained within the paper and edited the manuscript.

Funding statement. H.M. is supported by PAPIIT IN201214-25 and CONACYT 151223. J.A.G. is supported by the Natural Sciences and Engineering Research Council of Canada. L.J.T. is supported by grants from the Nature Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research. W.T.F. thanks the ERC (Advanced Grant SOMACCA #230604) for support. M.R. has been supported by the MIT Department of Linguistics and Philosophy as well as the Zukunftskonzept at TU Dresden funded by the Exzellenzinitiative of the Deutsche Forschungsgemeinschaft.

Competing interests. The authors have no competing interests.

References

- Honing H. 2013 Structure and interpretation of rhythm in music. In *Psychology of music* (ed. D Deutsch), 3rd edn, pp. 369–404. London, UK: Academic Press.
- Large EW, Palmer C. 2002 Perceiving temporal regularity in music. *Cogn. Sci.* **26**, 1–37. (doi:10.1207/s15516709cog2601_1)
- Large EW, Jones MR. 1999 The dynamics of attending: how people track time-varying events. *Psychol. Rev.* **106**, 119–159. (doi:10.1037/0033-295X.106.1.119)
- Drake C, Jones MR, Baruch C. 2000 The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition* **77**, 251–288. (doi:10.1016/S0010-0277(00)00106-2)
- Phillips-Silver J, Trainor LJ. 2007 Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition* **105**, 533–546. (doi:10.1016/j.cognition.2006.11.006)
- Winkler I, Haden G, Lading O, Sziller I, Honing H. 2008 Newborn infants detect the beat in music. *Proc. Natl Acad. Sci. USA* **106**, 2468–2471. (doi:10.1073/pnas.0809035106)
- 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)
- Zatorre RJ, Chen JL, Penhune VB. 2007 When the brain plays music. Auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558. (doi:10.1038/nrn2152)
- Phillips-Silver J, Trainor LJ. 2005 Feeling the beat in music: movement influences rhythm perception in infants. *Science* **308**, 1430. (doi:10.1126/science.1110922)
- Povel DJ, Essens P. 1985 Perception of temporal patterns. *Music Percept.* **2**, 411–440. (doi:10.2307/40285311)
- Yee W, Holleran S, Jones MR. 1994 Sensitivity to event timing in regular and irregular sequences: influences of musical skill. *Percept. Psychophys.* **56**, 461–471. (doi:10.3758/BF03206737)
- Teki S, Grube M, Kumar S, Griffiths TD. 2011 Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* **31**, 3805–3812. (doi:10.1523/JNEUROSCI.5561-10.2011)
- Merchant H, Battaglia-Mayer A, Georgopoulos AP. 2003 Interception of real and apparent circularly moving targets: psychophysics in human subjects and monkeys. *Exp. Brain Res.* **152**, 106–112. (doi:10.1007/s00221-003-1514-5)
- Mendez JC, Prado L, Mendoza G, Merchant H. 2011 Temporal and spatial categorization in human and non-human primates. *Front. Integr. Neurosci.* **5**, 50. (doi:10.3389/fnint.2011.00050)
- Repp BH, Su YH. 2013 Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* **20**, 403–452. (doi:10.3758/s13423-012-0371-2)
- Zarco W, Merchant H, Prado L, Mendez JC. 2009 Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* **102**, 3191–3202. (doi:10.1152/jn.00066.2009)
- Merchant H, Zarco W, Pérez O, Prado L, Bartolo R. 2011 Measuring time with different neural chronometers during a synchronization-continuation task. *Proc. Natl Acad. Sci. USA* **108**, 19 784–19 789. (doi:10.1073/pnas.1112933108)
- Donnet S, Bartolo R, Fernandes JM, Cunha JP, Prado L, Merchant H. 2014 Monkeys time their movement pauses and not their movement kinematics during a synchronization-continuation rhythmic task. *J. Neurophysiol.* **111**, 2138–2149. (doi:10.1152/jn.00802.2013)
- Doumas M, Wing AM. 2007 Timing and trajectory in rhythm production. *J. Exp. Psychol.* **33**, 442–455. (doi:10.1037/0096-1523.33.2.442)
- Konoike N, Mikami A, Miyachi S. 2012 The influence of tempo upon the rhythmic motor control in macaque monkeys. *Neurosci. Res.* **74**, 64–67. (doi:10.1016/j.neures.2012.06.002)
- Honing H, Merchant H. 2014 Differences in auditory timing between human and non-human primates. *Behav. Brain Sci.* **37**, 373–374.
- Patel AD. 2014 The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol.* **12**, e1001821. (doi:10.1371/journal.pbio.1001821)
- 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. (doi:10.1007/s00221-004-2159-8)
- Fitch WT, Rosenfeld AJ. 2007 Perception and production of syncopated rhythms. *Music Percept.* **25**, 43–58. (doi:10.1525/mp.2007.25.1.43)
- Repp BH, Iversen JR, Patel AD. 2008 Tracking an imposed beat within a metrical grid. *Music Percept.* **26**, 1–18. (doi:10.1525/mp.2008.26.1.1)
- Bengtsson SL, Ullén F, Henrik Ehrsson H, Hashimoto T, Kito T, Naito E, Forssberg H, Sadato N. 2009 Listening to rhythms activates motor and premotor cortices. *Cortex* **45**, 62–71. (doi:10.1016/j.cortex.2008.07.002)
- Chen JL, Penhune VB, Zatorre RJ. 2008 Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* **18**, 2844–2854. (doi:10.1093/cercor/bhn042)
- Chen JL, Penhune VB, Zatorre RJ. 2008 Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J. Cogn. Neurosci.* **20**, 226–239. (doi:10.1162/jocn.2008.20018)
- Chen JL, Zatorre RJ, Penhune VB. 2006 Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage* **32**, 1771–1781. (doi:10.1016/j.neuroimage.2006.04.207)
- Grahn JA, Brett M. 2007 Rhythm perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**, 893–906. (doi:10.1162/jocn.2007.19.5.893)
- Grahn JA. 2009 The role of the basal ganglia in beat perception: neuroimaging and neuropsychological investigations. *Ann. NY Acad. Sci.* **1169**, 35–45. (doi:10.1111/j.1749-6632.2009.04553.x)
- Grahn JA, Rowe JB. 2009 Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* **29**, 7540–7548. (doi:10.1523/JNEUROSCI.2018-08.2009)

33. Grahn JA, Rowe JB. 2013 Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. *Cereb. Cortex* **23**, 913–921. (doi:10.1093/cercor/bhs083)
34. Kung SJ, Chen JL, Zatorre RJ, Penhune VB. 2013 Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat. *J. Cogn. Neurosci.* **25**, 401–420. (doi:10.1162/jocn_a_00325)
35. Lewis PA, Wing AM, Pope PA, Praamstra P, Miall RC. 2004 Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia* **42**, 1301–1312. (doi:10.1016/j.neuropsychologia.2004.03.001)
36. Schubotz RI, Friederici AD, von Cramon DY. 2000 Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *NeuroImage* **11**, 1–12. (doi:10.1006/nimg.1999.0514)
37. Ullén F, Bengtsson SL. 2003 Independent processing of the temporal and ordinal structure of movement sequences. *J. Neurophysiol.* **90**, 3725–3735. (doi:10.1152/jn.00458.2003)
38. Vuust P, Ostergaard L, Pallesen KJ, Bailey C, Roepstorff A. 2009 Predictive coding of music—brain responses to rhythmic incongruity. *Cortex* **45**, 80–92. (doi:10.1016/j.cortex.2008.05.014)
39. Vuust P, Roepstorff A, Wallentin M, Mouridsen K, Ostergaard L. 2006 It don't mean a thing. . . : keeping the rhythm during polyrhythmic tension, activates language areas (BA47). *NeuroImage* **31**, 832–841. (doi:10.1016/j.neuroimage.2005.12.037)
40. Ramnani N, Passingham RE. 2001 Changes in the human brain during rhythm learning. *J. Cogn. Neurosci.* **13**, 952–966. (doi:10.1162/089892901753165863)
41. Grube M, Cooper FE, Chinnery PF, Griffiths TD. 2010 Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proc. Natl Acad. Sci. USA* **107**, 11 597–11 601. (doi:10.1073/pnas.0910473107)
42. Grube M, Lee KH, Griffiths TD, Barker AT, Woodruff PW. 2010 Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Front. Psychol.* **1**, 171. (doi:10.3389/fpsyg.2010.00171)
43. Geiser E, Notter M, Gabrieli JDE. 2012 A corticostriatal neural system enhances auditory perception through temporal context processing. *J. Neurosci.* **32**, 6177–6182. (doi:10.1523/JNEUROSCI.5153-11.2012)
44. Grahn JA, Henry MJ, McAuley JD. 2011 fMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *NeuroImage* **54**, 1231–1243. (doi:10.1016/j.neuroimage.2010.09.033)
45. Chapin HL, Zanto T, Jantzen KJ, Kelso SJA, Steinberg F, Large EW. 2010 Neural responses to complex auditory rhythms: the role of attending. *Front. Psychol.* **1**, 224. (doi:10.3389/fpsyg.2010.00224)
46. Riecker A, Kassubek J, Gröschel K, Grodd W, Ackermann H. 2006 The cerebral control of speech tempo: opposite relationship between speaking rate and BOLD signal changes at striatal and cerebellar structures. *NeuroImage* **29**, 46–53. (doi:10.1016/j.neuroimage.2005.03.046)
47. Riecker A, Wildgruber D, Mathiak K, Grodd W, Ackermann H. 2003 Parametric analysis of rate-dependent hemodynamic response functions of cortical and subcortical brain structures during auditorily cued finger tapping: a fMRI study. *NeuroImage* **18**, 731–739. (doi:10.1016/S1053-8119(03)00003-X)
48. Fraisse P. 1984 Perception and estimation of time. *Annu. Rev. Psychol.* **35**, 1–36. (doi:10.1146/annurev.ps.35.020184.000245)
49. van Noorden L, Moelants D. 1999 Resonance in the perception of musical pulse. *J. New Music Res.* **28**, 43–66. (doi:10.1076/jnmr.28.1.43.3122)
50. Chen Y, Repp BH, Patel AD. 2002 Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Hum. Mov. Sci.* **21**, 515–532. (doi:10.1016/S0167-9457(02)00138-0)
51. Grahn JA, Manly T. 2012 Common neural recruitment across diverse sustained attention tasks. *PLoS ONE* **7**, e49556. (doi:10.1371/journal.pone.0049556)
52. Hove MJ, Fairhurst MT, Kotz SA, Keller PE. 2013 Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *NeuroImage* **67**, 313–321. (doi:10.1016/j.neuroimage.2012.11.032)
53. McAuley JD, Henry MJ. 2010 Modality effects in rhythm processing: auditory encoding of visual rhythms is neither obligatory nor automatic. *Att. Percep. Psychophys* **72**, 1377–1389. (doi:10.3758/APP.72.5.1377)
54. Repp BH, Penel A. 2002 Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 1085–1099. (doi:10.1037/0096-1523.28.5.1085)
55. Merchant H, Zarco W, Prado L. 2008 Do we have a common mechanism for measuring time in the hundred of milliseconds range? Evidence from multiple interval timing tasks. *J. Neurophysiol.* **99**, 939–949. (doi:10.1152/jn.01225.2007)
56. Marchant JL, Driver J. 2013 Visual and audiovisual effects of isochronous timing on visual perception and brain activity. *Cereb. Cortex* **23**, 1290–1298. (doi:10.1093/cercor/bhs095)
57. Rao SM, Harrington DL, Haaland KT, Bobholz JA, Cox RW, Binder JR. 1997 Distributed neural systems underlying the timing of movements. *J. Neurosci.* **17**, 5528–5535.
58. Halsband U, Ito N, Freund HJ. 1993 The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain* **116**, 243. (doi:10.1093/brain/116.1.243)
59. Honing H, Bouwer F, Håden GP. 2014 Perceiving temporal regularity in music: the role of auditory event-related potentials (ERPs) in probing beat perception. In *Neurobiology of interval timing* (eds V Merchant, H de Lafuente). Berlin, Germany: Springer.
60. Honing H, Merchant H, Heden G, Prado LA, Bartolo R. 2012 Rhesus monkeys (*Macaca mulatta*) can detect rhythmic groups in music, but not the beat. *PLoS ONE*, **7**, e51369. (doi:10.1371/journal.pone.0051369)
61. Näätänen R, Paavilainen P, Rinne T, Alho K. 2007 The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* **118**, 2544–2590. (doi:10.1016/j.clinph.2007.04.026)
62. Selezneva E, Deike S, Knyazeva S, Scheich H, Brechmann A, Brosch M. 2013 Rhythm sensitivity in macaque monkeys. *Front. Syst. Neurosci.* **7**, 49. (doi:10.3389/fnsys.2013.00049)
63. 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**, 10 383–10 388. (doi:10.1073/PNAS.1402039111)
64. Marie C, Fujioka T, Herrington L, Trainor LJ. 2012 The high-voice superiority effect in polyphonic music is influenced by experience: a comparison of musicians who play soprano-range compared to bass-range instruments. *Psychomusicology* **22**, 97–104. (doi:10.1037/a0030858)
65. Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. 2005 Automatic encoding of polyphonic melodies in musicians and non-musicians. *J. Cogn. Neurosci.* **17**, 1578–1592. (doi:10.1162/089892905774597263)
66. Marie C, Trainor LJ. 2014 Early development of polyphonic sound encoding and the high voice superiority effect. *Neuropsychologia* **57**, 50–58. (doi:10.1016/j.neuropsychologia.2014.02.023)
67. Marie C, Trainor LJ. 2012 Development of simultaneous pitch encoding: infants show a high voice superiority effect. *Cereb. Cortex* **23**, 660–669. (doi:10.1093/cercor/bhs050)
68. Trainor LJ, Marie C, Bruce IC, Bidelman GM. 2014 Explaining the high-voice superiority effect in polyphonic music: evidence from cortical evoked potentials and peripheral auditory models. *Hear. Res.* **308**, 60–70. (doi:10.1016/j.heares.2013.07.014)
69. Arnal LH. 2012 Predicting 'When' using the motor system's beta-band oscillations. *Front. Hum. Neurosci.* **6**, 225. (doi:10.3389/fnhum.2012.00225)
70. Fujioka T, Trainor LJ, Large EW, Ross B. 2009 Beta and gamma rhythms in human auditory cortex during musical beat processing. *Ann. NY Acad. Sci.* **1169**, 89–92. (doi:10.1111/j.1749-6632.2009.04779.x)
71. 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. (doi:10.1523/JNEUROSCI.4107-11.2012)
72. 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. (doi:10.1126/science.1154735)

73. Arnal LH, Wyart V, Giraud AL. 2011 Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* **14**, 797–801. (doi:10.1038/nn.2810)
74. Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE. 2011 Tuning for the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* **31**, 3176–3185. (doi:10.1523/JNEUROSCI.4518-10.2011)
75. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P. 2010 Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* **20**, 172–176. (doi:10.1016/j.conb.2010.02.010)
76. Large EW. 2008 Resonating to musical rhythm: theory and experiment. In *Psychology of time* (ed. S Grondin), pp. 189–232. Bingley, UK: Emerald Group Publishing.
77. Nozaradan S, Peretz I, Missal M, Mouraux A. 2011 Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* **31**, 10 234–10 240. (doi:10.1523/JNEUROSCI.0411-11.2011)
78. Nozaradan S, Peretz I, Mouraux A. 2012 Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J. Neurosci.* **32**, 17 572–17 581. (doi:10.1523/JNEUROSCI.3203-12.2012)
79. 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. (doi:10.1016/j.neuron.2007.09.038)
80. Arnal LH, Doelling KB, Poeppel D. In press. Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* (doi:10.1093/cercor/bhu103)
81. Arnal LH, Giraud AL. 2012 Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **16**, 390–398. (doi:10.1016/j.tics.2012.05.003)
82. Schubotz RI. 2007 Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* **11**, 211–218. (doi:10.1016/j.tics.2007.02.006)
83. Tian X, Poeppel D. 2010 Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychol.* **1**, 166. (doi:10.3389/fpsyg.2010.00166)
84. Large EW, Snyder JS. 2009 Pulse and meter as neural resonance. The neurosciences and music III—disorders and plasticity. *Ann. NY Acad. Sci.* **1169**, 46–57. (doi:10.1111/j.1749-6632.2009.04550.x)
85. Iversen JR, Repp BH, Patel AD. 2009 Top-down control of rhythm perception modulates early auditory responses. *Ann. NY Acad. Sci.* **1169**, 58–73. (doi:10.1111/j.1749-6632.2009.04579.x)
86. Engel AK, Fries P. 2010 Beta-band oscillations—signalling the status quo? *Curr. Opin. Neurobiol.* **20**, 156–165. (doi:10.1016/j.conb.2010.02.015)
87. Gerloff C, Richard J, Hadley J, Schulman AE, Honda M, Hallett M. 1998 Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements. *Brain* **121**, 1513–1531. (doi:10.1093/brain/121.8.1513)
88. Pfurtscheller G. 1981 Central beta rhythm during sensorimotor activities in man. *Electroencephalogr. Clin. Neurophysiol.* **51**, 253–264. (doi:10.1016/0013-4694(81)90139-5)
89. Pollok B, Südmeyer M, Gross J, Schnitzler A. 2005 The oscillatory network of simple repetitive bimanual movements. *Brain Res. Cogn. Brain Res.* **25**, 300–311. (doi:10.1016/j.cogbrainres.2005.06.004)
90. Salmelin R, Hamalainen M, Kajola M, Hari R. 1995 Functional segregation of movement-related rhythmic activity in the human brain. *Neuroimage* **2**, 237–243. (doi:10.1006/nimg.1995.1031)
91. Toma K, Mima T, Matsuoka T, Gerloff C, Ohnishi T, Koshy B, Andres F, Hallett M. 2002 Movement rate effect on activation and functional coupling of motor cortical areas. *J. Neurophysiol.* **88**, 3377–3385. (doi:10.1152/jn.00281.2002)
92. Cirelli LK, Bosnyak D, Manning FC, Spinelli C, Marie C, Fujioka T, Ghahremani A, Trainor LJ. 2014 Beat-induced fluctuations in auditory cortical beta-band activity: using EEG to measure age-related changes. *Front. Psychol.* **5**, 742. (doi:10.3389/fpsyg.2014.00742)
93. Cravo AM, Rohenkohl G, Wyart V, Nobre AC. 2011 Endogenous modulation of low frequency oscillations by temporal expectations. *J. Neurophysiol.* **106**, 2964–2972. (doi:10.1152/jn.00157.2011)
94. Nelson A, Schneider DM, Takatoh J, Sakurai K, Wang F, Mooney R. 2013 A circuit for motor cortical modulation of auditory cortical activity. *J. Neurosci.* **33**, 14 342–14 353. (doi:10.1523/JNEUROSCI.2275-13.2013)
95. Hoppensteadt FC, Izhikevich EM. 1997 *Weakly connected neural networks*. New York, NY: Springer.
96. Mendoza G, Merchant H. 2014 Motor system evolution and the emergence of high cognitive functions. *Prog. Neurobiol.* **121**, 20–41. (doi:10.1016/j.pneurobio.2014.09.001)
97. Buzsáki G, Anastassiou CA, Koch C. 2012 The origin of extracellular fields and currents: EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* **13**, 407–420. (doi:10.1038/nrn3241)
98. Bartolo R, Prado L, Merchant H. 2014 Information processing in the primate basal ganglia during sensory guided and internally driven rhythmic tapping. *J. Neurosci.* **34**, 3910–3923. (doi:10.1523/JNEUROSCI.2679-13.2014)
99. Snyder JS, Large EW. 2005 Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cogn. Brain Res.* **24**, 117–126. (doi:10.1016/j.cogbrainres.2004.12.014)
100. Zanto TP, Large EW, Fuchs A, Kelso JAS. 2005 Gamma-band responses to perturbed auditory sequences: evidence for synchronization of perceptual processes. *Music Percept.* **22**, 531–547. (doi:10.1525/mp.2005.22.3.531)
101. Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE. 2009 The leading sense: supramodal control of neurophysiological context by attention. *Neuron* **64**, 419–430. (doi:10.1016/j.neuron.2009.10.014)
102. Schroeder CE, Lakatos P. 2009 Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* **32**, 9–18. (doi:10.1016/j.tins.2008.09.012)
103. Steriade M, Dossi RC, Nunez A. 1991 Network modulation of a slow intrinsic oscillation of cat thalamocortical neurons implicated in sleep delta waves: cortically induced synchronization and brainstem cholinergic suppression. *J. Neurosci.* **11**, 3200–3217.
104. Nácher V, Ledberg A, Deco G, Romo R. 2013 Coherent delta-band oscillations between cortical areas correlate with decision making. *Proc. Natl Acad. Sci. USA* **110**, 15 085–15 090. (doi:10.1073/pnas.1314681110)
105. Fries P. 2009 Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* **32**, 209–224. (doi:10.1146/annurev.neuro.051508.135603)
106. Kimura M. 1992 Behavioral modulation of sensory responses of primate putamen neurons. *Brain Res.* **578**, 204–214. (doi:10.1016/0006-8993(92)90249-9)
107. Merchant H, Zainos A, Hernández A, Salinas E, Romo R. 1997 Functional properties of primate putamen neurons during the categorization of tactile stimuli. *J. Neurophysiol.* **77**, 1132–1154.
108. Merchant H *et al.* 2015 Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas. *Adv. Exp. Med. Biol.* **829**, 185–196.
109. Merchant H, Perez O, Zarco W, Gamez J. 2013 Interval tuning in the primate medial premotor cortex as a general timing mechanism. *J. Neurosci.* **33**, 9082–9096. (doi:10.1523/JNEUROSCI.5513-12.2013)
110. Crowe DA, Zarco W, Bartolo R, Merchant H. 2014 Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J. Neurosci.* **34**, 11 972–11 983. (doi:10.1523/JNEUROSCI.2177-14.2014)
111. Merchant H, Harrington DL, Meck WH. 2013 Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* **36**, 313–336. (doi:10.1146/annurev-neuro-062012-170349)
112. Perez O, Kass RE, Merchant H. 2013 Trial time warping to discriminate stimulus-related from movement-related neural activity. *J. Neurosci. Methods* **212**, 203–210. (doi:10.1016/j.jneumeth.2012.10.019)
113. Nagarajan SS, Blake DT, Wright BA, Byl N, Merzenich M. 1998 Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *J. Neurosci.* **18**, 1559–1570.
114. Bartolo R, Merchant H. 2009 Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. *Exp. Brain Res.* **197**, 91–100. (doi:10.1007/s00221-009-1895-1)
115. Merchant H, de Lafuente V, Peña F, Larriva-Sahd J. 2012 Functional impact of interneuronal inhibition in the cerebral cortex of behaving animals. *Prog.*

- Neurobiol.* **99**, 163–178. (doi:10.1016/j.pneurobio.2012.08.005)
116. Gewaltig MO, Diesmann M, Aertsen A. 2001 Propagation of cortical synfire activity: survival probability in single trials and stability in the mean. *Neural Netw.* **14**, 657–673. (doi:10.1016/S0893-6080(01)00070-3)
 117. Nagasaka Y, Chao ZC, Hasegawa N, Notoya T, Fujii N. 2013 Spontaneous synchronization of arm motion between Japanese macaques. *Sci. Rep.* **3**, 1151. (doi:10.1038/srep01151)
 118. Guttman SE, Gilroy LA, Blake R. 2005 Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychol. Sci.* **16**, 228–235. (doi:10.1111/j.0956-7976.2005.00808)
 119. Brodsky W, Henik A, Rubinstein BS, Zorman M. 2003 Auditory imagery from musical notation in expert musicians. *Percept. Psychophys.* **65**, 602–612. (doi:10.3758/BF03194586)
 120. Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TEJ. 2008 The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* **11**, 426–428. (doi:10.1038/nn2072)
 121. Merchant H, Perez O, Bartolo R, Mendez JC, Mendoza G, Gamez J, Yc K, Prado L. In press. Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *Eur. J. Neurosci.*
 122. Hove MJ, Iversen JR, Zhang A, Repp BH. 2013 Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychol. Res.* **77**, 388–398. (doi:10.1007/s00426-012-0441-0)
 123. Marr D. 1982 *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W H Freeman & Co.
 124. Lerdahl F, Jackendoff R. 1983 *A generative theory of tonal music*. Cambridge, MA: MIT Press.
 125. Longuet-Higgins HC, Lee CS. 1982 The perception of musical rhythms. *Perception* **11**, 115–128. (doi:10.1068/p110115)
 126. Desain P, Honing H. 1999 Computational models of beat induction: the rule-based approach. *J. New Music Res.* **28**, 29–42. (doi:10.1076/jnmr.28.1.29.3123)
 127. Dixon S. 2007 Evaluation of the audio beat tracking system BeatRoot. *J. New Music Res.* **36**, 39–50. (doi:10.1080/09298210701653310)
 128. Gouyon F, Dixon S. 2005 A review of automatic rhythm description systems. *Comput. Music J.* **29**, 34–54. (doi:10.1162/comj.2005.29.1.34)
 129. Large EW, Kolen JF. 1994 Resonance and the perception of musical meter. *Connect. Sci.* **6**, 177–208. (doi:10.1080/09540099408915723)
 130. Velasco MJ, Large EW. 2010 Entrainment to complex rhythms: tests of a neural model. In *Abstracts of Neuroscience 2010, 13–17 November 2010, San Diego, CA*, 324.3. Society for Neuroscience.
 131. Snyder J, Krumhansl CL. 2001 Tapping to ragtime: cues to pulse finding. *Music Percept.* **18**, 455–489. (doi:10.1525/mp.2001.18.4.455)
 132. Hattori Y, Tomonaga M, Matsuzawa T. 2013 Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Sci. Rep.* **3**, 1566. (doi:10.1038/srep01566)
 133. Cook P, Rouse A, Wilson M, Reichmuth CJ. 2013 A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* **127**, 1–16. (doi:10.1037/a0032345)
 134. Hasegawa A, Okanoya K, Hasegawa T, Seki Y. 2011 Rhythmic synchronization tapping to an audio–visual metronome in budgerigars. *Sci. Rep.* **1**, 120. (doi:10.1038/srep00120)
 135. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. (doi:10.1016/j.cub.2009.03.038)
 136. Schachner A. 2010 Auditory-motor entrainment in vocal mimicking species: additional ontogenetic and phylogenetic factors. *Commun. Integr. Biol.* **3**, 290–293. (doi:10.4161/cib.3.3.11708)
 137. Ravignani A, Bowling D, Fitch WT. 2014 Chorusing, synchrony and the evolutionary functions of rhythm. *Front. Psychol.* **5**, 118. (doi:10.3389/fpsyg.2014.01118)
 138. Ackermann H, Hage S, Ziegler W. In press. Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav Brain Sci.* **37**, 529–546.
 139. 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. (doi:10.3389/fnsys.2014.00057)