HAL open science

Psychological and neuroscientific foundations of rhythms and timing

Keith Doelling, Sophie Herbst, Luc Arnal, Virginie van Wassenhove

To cite this version:

Keith Doelling, Sophie Herbst, Luc Arnal, Virginie van Wassenhove. Psychological and neuroscientific foundations of rhythms and timing. 2023. pasteur-03924789

HAL Id: pasteur-03924789

https://pasteur.hal.science/pasteur-03924789v1

Preprint submitted on 5 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

This is a draft of a Chapter that has been accepted for publication by Oxford University Press in the forthcoming book Performing Time edited by Clemens Wöllner & Justin London due for publication in 2022.

Psychological and neuroscientific foundations of rhythms and timing

[1] Keith Doelling, [2] Sophie Herbst, [1] Luc Arnal, [2] Virginie van Wassenhove

[1] Institut de l'Audition, Institut Pasteur, Paris, France. [2] Neurospin, Cognitive Neuroimaging Unit, CEA, INSERM, CNRS, Université Paris-Saclay, France.

Abstract
Cultural artefacts, such as dance and music, are characterized by temporal properties, broadly referred to as *rhythms.* The temporal structures that emerge when individuals synchronize their movements provide a sense of togetherness and common fate, even if individuals can readily tune in and out of this shared temporal space. In this chapter, we concisely discuss the endogenous and exogenous factors contributing to the emergence of brain rhythms and rhythmic behaviors, and how their interplay contributes to sophisticated forms of expressions in humans. The capacity for coupling and uncoupling voluntarily from external rhythms creates a potential tension between our environment (exogenously driven) and our internal state (endogenously driven) that can be exploited for artistic benefit, in the form of *surprisal*.

Introduction

Rhythms are everywhere: in the motion of planets that sets the pace of our day/night routines, in the rhymes that humans enjoy reading, in the music and dance they generate. At each scale of life, from molecules to flocks of birds, rhythms mark time and provide the metrics for information flow. In György Ligeti's Poème Symphonique (1962), a hundred metronomes, each set to a distinct tempo, are launched as simultaneously as possible, marking time for a few minutes. This symphonic poem could be thought of as an idealized metaphor of brain rhythms: hundreds of neural populations can be rhythmically active at the same time, at the same or at different frequencies, in phase or out-of-phase with each other. The multiple time metrics, endogenous to brain function, may serve the coding, segmentation, regulation, and transmission of information. This chapter provides a broad neuroscientific view of the psychological and neural constraints of rhythms and rhythmic processing, leaving the subtleties of individual areas of expertise to the other chapters in this book. We first discuss the importance and definition of rhythms in human productions, then turn to the role of neural oscillations, illustrating specific roles of rhythms for prediction, attention, and anticipation—notions that are central to artistic productions. Finally, we highlight a tension inherent to biology and psychology, namely, the interfacing between exogenous temporalities and endogenous bodily rhythms that make individuals' clocks relative.

1 Rhythms

In this chapter, rhythms are defined as periodic patterns in signals (e.g., sounds, body movements, or neural dynamics) over a wide range of time scales. Temporal patterns do not have to be strictly isochronous to qualify as rhythm; indeed, these rhythms can be very complex, such as the hierarchically nested structures in music, dance, or speech that humans readily produce. For this chapter, our case study will mostly be a quasi-isochronous single stream of

sounds (i.e., notes, syllables, or anything else). The accompanying constructs representing higher order temporal abstractions such as beat, tempo, and meter are thoroughly described in other chapters (e.g. London; Henry & Kotz; Hammerschmidt, this volume). Herein, we focus on the sequence itself, which unfolds chronologically, with the expectation of when the next note or gesture will occur given the preceding temporal context. To keep the scope of this chapter open to a variety of research areas, our focus will remain at this low-level definition of rhythm.

1.1 Human production of rhythms

Across cultures and eras, human activities and the artefacts generated by them, which we refer to as human productions, have been characterized by time constants that occur and re-occur as we create new works of art or complete menial tasks. For example, handwriting, a behavior not typically thought of as rhythmic, can be modeled as a coupled oscillator system, as if the pen were being pulled by two pendulums swinging in perpendicular directions (Gangadhar et al., 2007; Hollerbach, 1981). Similar rhythmicity can be found in sweeping, painting, walking and swimming, and of course, in speech, dance, and music. These rhythmic sequences of behaviors tend to occur any time repetition is required. When a single act is not enough to accomplish the task, repeating it allows us to build on past actions, creating more complex structures. In performing similar acts repeatedly, we necessarily create rhythms.

At first inspection, a behavior might only be considered rhythmic when the same temporal interval, or an exact integer multiple of it, is repeated precisely over a sequence. A metronome might come to mind as a possible example of a prototypical isochronous rhythm being used to dictate the tempo of a musical piece. Most human behaviors would fail by this definition, even ones that we would expect to succeed. Walking, for example, an undeniably rhythmic action, shows fractal variance over time (Hausdorff et al., 1996). Instead, by our definition, rhythms are sequences with temporal structures, whereby each event provides some – but not necessarily exact – information about when the next event will occur: a balance between predictability and complexity.

To illustrate the point further, consider audition, in which rhythmic human production leads to rhythmic sensory stimuli. In speech and music, each unit of action (e.g., articulating a syllable or singing a note) is similar to each other unit—for example, in the opening and closing of the vocal tract - as they occur within similar temporal constraints. Still, despite their similarity, each unit is intentionally distinguished from one another to form the building blocks of complex concepts for listeners. These differences create variety not only in the actions generated, but also in the time it takes to execute them. These signals are therefore rhythmic, in the sense that they maintain temporal regularities and structure without being isochronous. Furthermore, the timing of speech, dance, or music is not passive; instead, it actively nuances communication. Consider the slowing down of the syllabic rate to indicate the end of a thought (Local et al., 1986), or the expressive timing of a musician to give a performance greater feeling (Clarke, 1989; Todd, 1985). Herein, “rhythms” refer to repetitive temporal structures that do not precisely match isochrony although they rely on periodic processes. While human productions like speech, music, or dance are sourced from the rhythmic tools provided to us by our biology, we must also appreciate the complexity of our timing, which is added to accommodate the semantic and emotional processes that these signals are intended for.

Evidence for rhythmicity in human productions comes often from corpus analyses, which analyze large databases of audio signals to investigate their temporal properties. Using the Fourier Transform (Figure 1), we can glean whether a signal contains high or low frequency content. The presence of rhythmicity depends on finding a peak in the spectrum, noting for example that there is a dominance of power in the 5 Hz range i.e., at a periodicity of 200 milliseconds (ms). The width of the spectral peak will vary with how precise or isochronous a

rhythm is. For instance, in Ligeti's Poème Symphonique the rhythmic complexity is apparent in the number of spectral peaks that also overlap (Figure 1A) or as *"spiderwebs of rhythm emerg[ing] from the cloud of ticks"* as Alex Ross put it (2007, pp. 508). Ding and colleagues (2017) used this method to investigate a wide array of recordings for speech and music. The languages they analyzed showed an average peak frequency of 5 Hz with a variance of about ±1 Hz and the music they analyzed, an average peak around 2 Hz (a periodicity of 500 ms) across a wide array of instruments (single or ensemble pieces) and musical styles. The authors suggest that the observed spectral peaks may reflect the beat rate with their metrical subdivisions, such as those typically reported in Western music (London, 2012) and readily tracked by cortical activity (Nozaradan et al., 2011, 2012, discussed below). While the presence of rhythm in music is uncontroversial, how these rhythms compare across cultures is less clear. Recent work (Jacoby & McDermott, 2017; Polak et al., 2018) has taken an innovative, iterative tapping approach to find which rhythmic patterns are endemic to particular cultures. They found that while each culture had a different pattern seemingly typical of their own culture (e.g., Balkan Folk Music contains 3 - 3 - 2, a pattern that was more heavily represented in people from that region), they also reported a universal preference for rhythms with low integer ratio relationships between them.

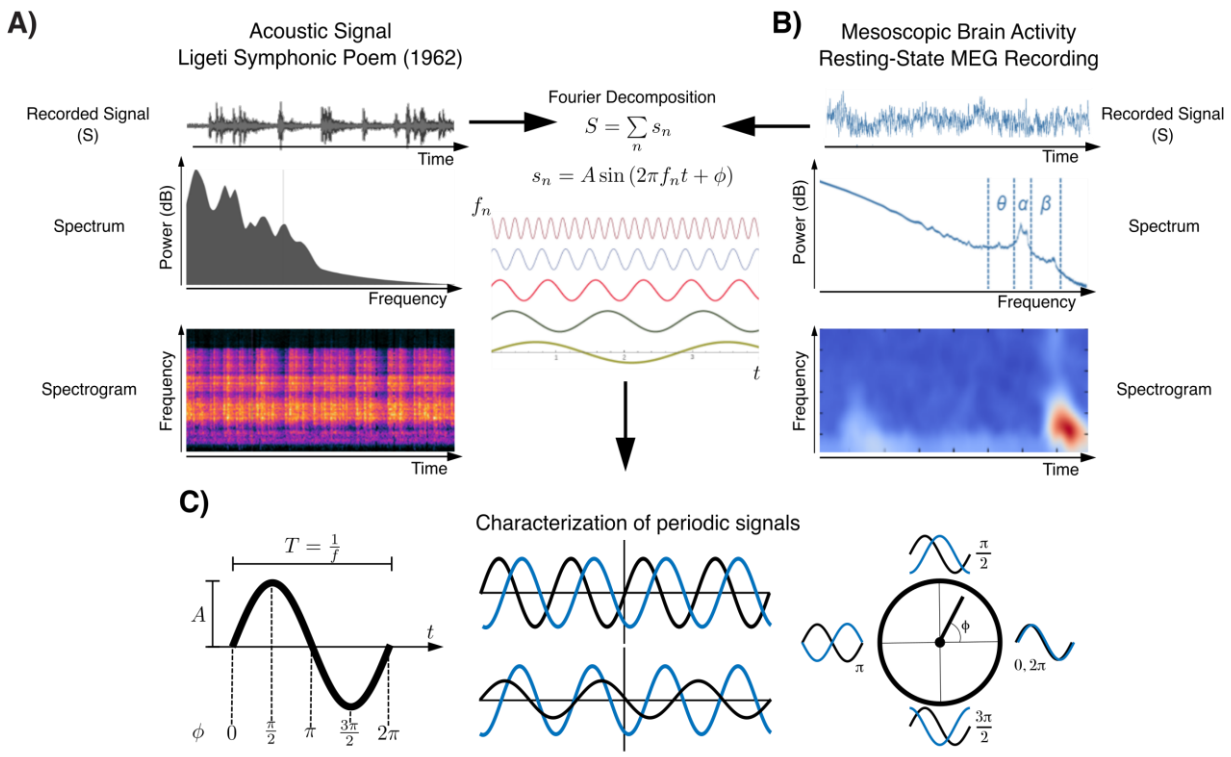

Figure 1: Signal processing of time series. **A**) Example of an audio recording of Ligeti's Poème Symphonique for 100 metronomes (left) and **B**) a resting-state data acquired with magnetoencephalography (MEG). The *Fourier decomposition* provides a quantitative description of signals as a sum of sinusoids. The power spectral density or spectrum is characterized by a 1/f spectrum in all physical systems so that the lower frequencies display a higher power than higher frequencies. Peaks in the spectrum indicate the most prominent periodicities. In Ligeti's spectrum, each peak is a metronome. In the brain's spectrum, alpha band (α, 8-12 Hz) is a spontaneous neural oscillation seen during resting-state. We also see peaks of activity at 4-7 Hz in the theta (θ) band, and 20-30 Hz in the beta (β) band. Time-frequency analyses result in spectrograms preserving the temporal dimension of spectral fluctuations, which a spectral representation dismisses. The extreme regularities of Ligeti's symphonic poem become evident (bottom left panel; yellow pattern) in contrast to the abrupt increases in a specific part of the

spectrum in brain activity (bottom right panel, red). **C**) The outcome of Fourier decomposition helps characterizing the periodic components of signals by their frequencies (f), period (T), amplitudes (A) and relative phases (ϕ), (left panel). The middle panel illustrates two sinusoids with the same frequency but different phases (top) or with different frequencies and amplitudes (bottom). The full range of phase differences between two signals sharing the same frequency or period is illustrated on the unit circle (right panel). The in-phase and out-of-phase relation of the two sinusoids (0 and π, respectively) illustrate that sharing the same frequency does not equate synchrony or shared temporal alignment.

The preference for rhythmicity in speech and music (and likely dance) results from temporal constraints of at least two types: the temporal constraints of our body effectors and those of our neural architecture. With regards to the effectors, successful musical performance requires the manipulation of a wide range of articulators to dynamically control the timing and quality of the final outcome. For example, playing the piano requires both the fine-grained movement of the fingers as well as more large-scale arm, shoulder and core movements. The requirement for coordination between our articulators creates temporal constraints on its movement. There are limits to how fast and how slow we can reliably control these movements. On the other hand, the nature of our neural architecture requires time for processing as well. An artificially highly sped-up melody will create the perception of blurred individual notes and groupings not present at the intended tempo. To borrow an example from language, while both speech and sign languages are produced by vastly different effectors, their units of information move at nearly the same speed (Wilbur & Nolkn, 1986). That such remarkably different effectors have similar time constants suggests these actions are constrained by a third component, the neural time constraints of the listener.

As members of the same species, humans share similar neural and effector constraints. Our ability to enjoy and generate music that others can contribute to depends on creating a stimulus that behaves in a manner that is coherent with these shared constraints. It is possible then that these shared temporal properties can help explain the universal features of rhythmicity at the cultural level; still, why do humans exhibit a universal tendency to synchronize and embody auditory rhythms, as they do, for example, in dance?

1.2 Embodied time in a predictive brain

Rhythmicity yields the capacity to predict the timing of future events, which in turn has the potential to enhance perception and facilitate action. Internalizing and forming abstract representations of external temporal regularities (music or seeing others dancing), confers considerable evolutionary advantages by enabling adaptive and proactive behaviors. The temporal scales of spontaneous human productions are distributed around tens to hundreds of milliseconds (see also Hammerschmidt, this volume). Many phenomena related to perception and action occur at this scale, which likely accounts for the human propensity to synchronize movements to rhythms (Patel & Iversen, 2014).

At these time scales, the motor system is often recruited during the processing of rhythmic sequences, and it aligns its ongoing activity to anticipated and unattended events (Arnal & Giraud, 2012; Fujioka et al., 2009, 2012). This predictive alignment increases with sensorimotor expertise (Doelling & Poeppel, 2015) and improves with periodic motor priming (vocally) or auditory motor training (Cason et al., 2015). The relationship between the motor system and the perceptual processing of rhythmic sequences supports behavioral findings that temporal prediction in perception and action might rely on similar systems (Wöllner & Cañal-Bruland, 2010), and computational/neurophysiological mechanisms (Schubotz, 2007).

Whether the recruitment of the motor system is necessary for perception or merely useful in noisy or unfavorable listening or viewing conditions remains a matter of debate (Stokes et al., 2019). Many cerebral regions, for instance, the cerebellum (Bareš et al., 2019) can internalize implicit or explicit knowledge about the timing of sensory environments. This can arguably contribute to the anticipation of upcoming sensorimotor events, as well as control temporal predictions and expectations in sensory regions. It is also noteworthy that sensorimotor feedback loops create rhythmic behaviors even in the absence of external signals: postural control is generally viewed as a feedback loop in which the alternation between the sensory detection of the body's current position (proprioception) and its stabilization by motor correction creates perceptible cyclic behaviors (Massion, 1994).

In fact, the motor system can exploit a rich repertoire of actions (head nodding, tapping with the feet, dancing) in which the associated temporal trajectories of effectors can be internally emulated in the temporal dynamics of brain activity to predict the sensory consequences of actions. In doing so, the motor system uses top-down signals (i.e., volition) to tune the ongoing activity (and, specifically, the phase of ongoing oscillations) in sensory regions. This top-down control of the motor system over sensory activity enables the fine-tuning and temporal calibration of sensory regions to the impending sensory consequences. Owing to its expertise in exerting motor trajectories with sensorimotor control at the millisecond timescale, the motor system is particularly well-suited for time-related anticipatory computations, including the estimation of one's timing errors (Kononowicz et al., 2019). In this sense, the performance of dance (the realization of these motor trajectories) may serve as a physical instantiation of our own internal anticipation of time.

The experimental evidence showing that the motor system automatically synchronizes with external temporal patterns (periodic or aperiodic) suggests that top-down signals proactively track temporal patterns to facilitate their processing. In this view, temporal predictions in the auditory domain correspond to a covert form of active inference (Adams et al., 2013) in which top-down signals that are generated to dynamically control the production of audible actions are also generated during the perception of predictable auditory streams. By comparing simulated and real sensory outcomes, prediction errors can be computed, and used to learn, correct, and improve temporal predictions in the system.

Anticipation in time is essential to theories of brain function which propose that the brain generates internal models of the sensory world to predict the future (Bar, 2011). Brain theories like *predictive coding* or *active inference* provide a unified perspective on perception and action (Adams et al., 2013; Brown et al., 2011; Friston, 2005, 2018). Internal models use the same computational strategies to predict external events and the sensory consequences of our own actions, suggesting that the brain exploits all available sources of internal knowledge to reduce external and internal noise. The predictive and anticipatory exploration of the environment often requires several repetitions of basic movements (visual saccades, sniffing) to accumulate sufficient evidence and capture the full picture of a scene or of a scent. This sensory exploration, or *active sensing* (Schroeder et al., 2010), does not only occur in modalities that possess moveable motor effectors; it is also seen in audition and in somatosensation. Although humans cannot move their ears, exploring an acoustic scene is not a passive continuous process, but a discrete and repetitive one involving multiple looks in time (Viemeister & Wakefield, 1991). Empirical support for active sensing and the sampling of sensory evidence in the brain can be found in rhythmic patterns of neural activity, often referred to as neural oscillations.

1.3 Brain rhythms in perception, action, and cognition

Neural oscillations and oscillatory dynamics are a prominent feature of neural activity (Buzsáki, 2006; Wang, 2010). In the last decades, the field of neuroscience has shifted its opinion on the utility of neural oscillations: from epiphenomenal "fumes of computation" to functional implementation of binding, predicting, and orchestrating neural information, both locally and globally, in the networks (*e.g.* the "communication-through-coherence hypothesis"; Engel et al., 2001).

The origins of rhythmic and oscillatory neural activity depend on the scale of observation (Wang, 2010). Microscopic membrane fluctuations and the intrinsic properties of single neurons display resonant properties so that neuronal spikes are elicited at a preferred frequency (Llinás, 1988; Whittingstall & Logothetis, 2009). At the mesoscopic scale, neural oscillations reflect the mean activity of local neural assemblies; at the macroscopic scale, rhythmic fluctuations capture inter-areal synchronizations over time and space (Varela et al., 2001; Wang, 2010). Brain rhythms reflect the tendency of single neurons (and groups of neurons) to act as self-sustained oscillators, resonating in restricted frequency ranges. Neural oscillations display specific temporal periodicities (Figure 1B): delta (δ) occurs at 1-3 Hz, theta (θ) at 4 -7 Hz (i.e., periodicity of about 250 ms), and the most prominent alpha (α) at 7-12 Hz (i.e., periodicity of about 100 ms). Strikingly, these characteristic temporal scales or frequencies are preserved across species—independent, even, of brain sizes (Buzsáki et al., 2013). Therefore, these temporal scales reflect biophysically relevant time scales for behaviors across species, which may provide universal constants for cognition.

Neural oscillations can be identified in the spectral domain through a peak amplitude emerging at a restricted frequency (Donoghue et al., 2020). The identification often uses the Fourier Transform (Figure 1), which assumes a sinusoidal shape of the underlying neural signals. This is a strictly signal processing approach, one that can be readily applied to any time-varying signals (Figure 1A: Ligeti's symphonic poem; Figure 1B: an individual's brain activity in quiet wakefulness). The "Fourier fallacy" (Jasper, 1948) over-simplifies the characterizations of neural rhythms by neglecting their morphology (Cole & Voytek, 2017). This concern has prompted novel approaches to assess oscillatory activity more accurately, for instance by trying to preserve the dynamic spectral properties of the signals (Cole & Voytek, 2019) or using a statistical modeling approach (Dupré la Tour et al., 2017). Rhythmic activity depicted as a sinusoid is largely acknowledged as an oversimplification, although it is useful for conveying basic ideas.

Neural oscillations provide the temporal dimensions of how the mind works (Jones, 1976; Pöppel, 1971; Varela, 1999) and rely on a parsimonious account of brain functions. Neural oscillations and recurrent neural activity provide natural temporal structures with which information can be sampled in (perception) and out (action), likely forming a structural basis of cognition in the time domain. In the context of the predictive brain theories discussed above, neural oscillations embody the temporal metrics of predictions. In her early theoretical formalization of attentional processes, Jones (1976) described attention as a hierarchy of temporal metrics (rhythms) so that the loci of attention in time could occur at multiple time scales. The Dynamic Attending Theory, or DAT (Jones, 1976) was proposed to be implemented through neural oscillations, to account for sampling and the postulated "attentional pulses" (Large & Jones, 1999; Figure 2A). DAT postulates that the temporal structure of sensory inputs is internalized by endogenous rhythms, to improve sensory analysis and facilitate reaction times at rhythmically expected time points (Rimmele et al., 2018). In agreement with the rhythmic attentional regulation proposed in the DAT, behavioral performance has been shown to fluctuate in discrete, rhythmic sampling patterns, rather than according to a continuous attentional spotlight (Fiebelkorn et al., 2013; Ho et al., 2017; Landau & Fries, 2012).

The facilitation of processing rhythmically predicted inputs is rendered possible by neurophysiology: neuronal spiking activity tends to occur at preferred moments (or phases of

oscillatory activity) so that attentional or sensorimotor alignment of ongoing oscillations in time regulate the encoding of sensory information. This mechanism consists mainly in low-frequency neural oscillations (up to about 14 Hz) which synchronize to phasically defining windows of opportunity for neuronal spiking (Buzsáki, 2006; Panzeri et al., 2010; Wang, 2010; Figure 2B). In the following section, we will more thoroughly discuss the phase alignment of neural oscillations to external stimulation optimizes the encoding of sensory information (Lakatos et al., 2008; Panzeri et al., 2010; Schroeder et al., 2010; Whittingstall & Logothetis, 2009).

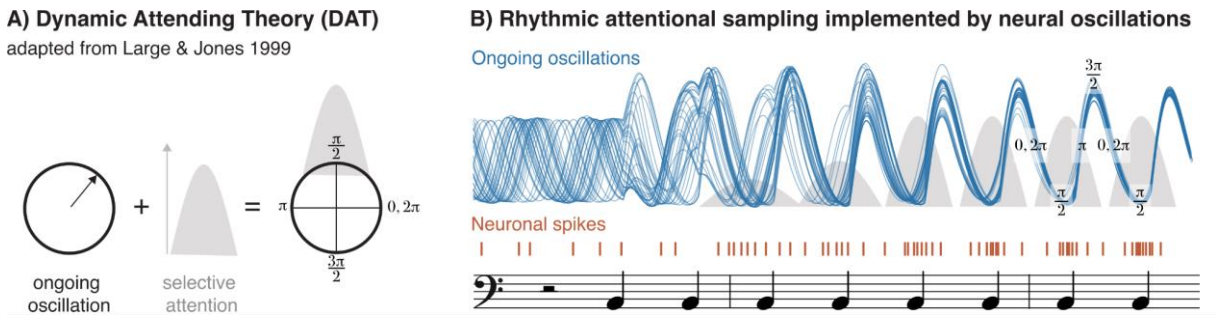

Figure 2: Rhythmic sampling. **A)** Dynamic Attending Theory (DAT), adapted from Large & Jones (1999). Selective attention (depicted by the gray cone) is aligned to a particular phase of an ongoing oscillation represented here by the phase cycle. The frequency of the oscillation defines the sampling rate, the phase aligns with the moment at which attention is highest. **B)** Simulating an implementation of rhythmic attention by oscillatory dynamics. Bottom, a rhythmic sequence of notes that stimulates oscillatory activity. Top, examples of ongoing oscillations as simulated by a Wilson-Cowan Model (Wilson & Cowan, 1972). Before the notes begin, oscillatory activity is at a random phase (rest period). The rhythmic stimulation elicits a series of phase shifts, which yields convergence to a specific phase over time, allowing for attention to unfold at the optimal phase. Middle, neuronal spikes biased by the excitability of the oscillation. Spikes are initially random but become locked to an optimal phase when the oscillation is synchronized.

2 Prediction, attention, and anticipation in time

Temporal predictions can be conveyed by (quasi-) isochronous rhythms (e.g., a musical beat occurring every 0.5 seconds), or isolated time intervals that tend to recur (e.g., the time from pushing the start button on your computer to opening your inbox). In the context of periodic sequences, the dynamics of cortical oscillations at a low sensory level represent a powerful means to extract and exploit temporal regularities. By modulating the allocation of processing resources in time, intrinsic oscillators play a functional role in temporal predictions and act as hard-wired temporal priors, periodically aligning neural excitability with the occurrence of repetitive sensorimotor inputs. DAT proposes that attention allocates limited resources efficiently to the most critical moments in time, selectively enhancing the processing of relevant events (Figure 2): the interplay between isochronous sensory inputs and attentional allocation in time may be aligned such that the attended moment (attention-driven) and the moment of temporal predictions (temporal statistics) are fully congruent, or in-phase.

Prediction (prior knowledge) and attention (selective relevance) are theoretically distinct processes (Summerfield & Egner, 2009; Wyart et al., 2012), but their dissociation must be clarified for the temporal domain, where, depending on task context, expectations might reflect both attentional orienting and sensory prediction (Doherty et al., 2005; Todorovic et al., 2015).

In the case of rhythms, the two concepts are related but refer to different aspects of the temporal process (de Lange et al., 2018; Press et al., 2020): attention increases the gain of neural responses for relevant timepoints, prediction exploits prior event times to assess when those timepoints might be, reducing the costs of processing incoming inputs (Arnal et al., 2015; Arnal & Giraud, 2012; Cravo et al., 2011).

As an illustration, both aspects are necessary when a conductor incrementally changes tempo. The orchestra or the corps de ballet must both predictively entrain their movements to the new rhythm and attend to the critical timepoints, beats, to optimize their perception. Section 2.1. discusses this specific scenario. Section 2.2. will then introduce the more complex consequences of the coexistence between periodic and aperiodic temporal predictions.

2.1 Entrainment: temporal predictions and attention in time in unison

Neural entrainment refers to a mechanism by which periodic sensory inputs induce the adaptation of a period, and the alignment of the phase of ongoing neural oscillations. In Figure 2B, the sequence of musical notes serves as a rhythmic input entraining ongoing neuronal oscillations; a simulated oscillator thereby develops synchrony and, in turn, regulates neuronal excitability to align to the rhythmic stimulus. Oscillatory entrainment has been proposed to be instrumental in temporal processing and sensory selection (Lakatos et al., 2008, 2019; Obleser & Kayser, 2019) by maintaining a temporal reference frame between sensory stimuli and neural responses (Herrmann et al., 2013).

Many studies have emphasized that the intrinsic auditory brain rhythms (notably, δ and θ; Figure 1B) match, to a remarkable degree, those found in speech or music; and there is ample experimental evidence that music and speech induce neural entrainment (Arnal & Giraud, 2012; Doelling et al., 2019; Luo & Poeppel, 2007). We have briefly mentioned the importance of sensorimotor alignment in action (Section 1.2) but neural entrainment also facilitates the processing of sensory inputs, both within and across sensory modalities (Besle et al., 2011; Busch & VanRullen, 2010; Henry & Obleser, 2012; Kösem et al., 2014; Stefanics et al., 2010). For instance, the behavioral detection of short silences while listening to sound fluctuates in phase with the external rhythm and with the measured neural oscillations at that same frequency range (Henry & Obleser, 2012). Visual sensitivity is also modulated in phase with rhythmic inputs (Cravo et al., 2013).

When rhythmic inputs are provided in a different sensory modality, both speed and behavioral performance in experimental tasks can improve. For instance, in a visual conjunction search paradigm, where participants had to detect a horizontal or a vertical bar that changed color rhythmically, the addition of sounds synchronized to the color change sped up participants' reaction times (Van der Burg et al., 2008). On the other hand, sounds synchronized to visual distractors both slowed participants down and decreased their accuracy (Kösem & Van Wassenhove, 2012; see also Miller et al., 2013). These behavioral results were interpreted in the context of the cross-sensory entrainment hypothesis (Lakatos et al., 2008), whereby a stream of events in one sensory modality (e.g., audition) can serve as a temporal structure for the other sensory stream (e.g. vision) so long as they share the same rhythm or frequency.

Following the observation of behavioral benefits of entrainment, an important working hypothesis is that entrainment flexibly synchronizes, and implicitly regulates, neuronal excitability for efficient processing of environmental inputs, be they auditory, visual, speech-related, musical, gestural, or multisensory (Arnal & Giraud, 2012; Giraud & Poeppel, 2012; Lakatos et al., 2008; Schroeder et al., 2010; Schroeder & Lakatos, 2009). Note that this working hypothesis entails the *a priori* existence of latent neural oscillators that have the propensity to

oscillate at a particular frequency. Although some oscillations are readily seen in resting-state (Figure 1B) this is difficult to establish with full certainty, and recent reconsiderations of the criteria for neural entrainment have been discussed both empirically and theoretically (Doelling & Assaneo, 2021; Lakatos et al., 2019; Obleser & Kayser, 2019). This point will be further discussed in Section 3.

2.2 Non-rhythmic temporal predictions

In the next section, we consider temporal predictions outside of a rhythmic context. Dancers must be sensitive not only to the rhythm when the piece is played but also to isolated temporal intervals throughout a performance, such as the time from when the conductor arrives on stage and when they will give the first cue, or how long a solo musician will hold a fermata. Such predictions are often memory-based, developed over the course of rehearsal. Furthermore, many musical styles employ non-periodic meters (Huron, 2008) that can nevertheless be anticipated and embodied in dance (Godøy et al., 2016; Su, 2016).

The notion of temporal predictions is sometimes conflated with that of entrainment to periodic sensory inputs; but, as we have seen, aperiodic temporal predictions can also occur in human productions (Rimmele et al., 2018). Recent work suggests that behavioral performance is equally enhanced for predictable conditions – whether periodic or not – compared to unpredictable ones (Bouwer et al., 2020; Morillon et al., 2016). This could present a challenge for classical entrainment models, which hypothesize that periodicity should maximize neural entrainment and ensuing behaviors. While potentially useful, entrainment is not the sole neural mechanism involved in temporal prediction.

Implicit predictions of temporal intervals have been studied using the *foreperiod paradigm* (Niemi & Näätänen, 1981; Woodrow, 1914): observers learn to predict from a cue stimulus the moment of appearance of a target stimulus, i.e. the foreperiod. Depending on the underlying temporal statistics, reaction times decrease at the most probable foreperiods (fixed foreperiod designs), but also for longer foreperiods (variable foreperiod designs) through conditional updating of the temporal prediction. A temporal prediction of the foreperiod has to be initiated anew at each presentation of the cue, which itself cannot be predicted in time, such that no isochronous temporal structure emerges. Importantly, human observers and non-human primates can learn and represent several foreperiod distributions concurrently or in succession (Bueti et al., 2010; Herbst et al., 2018; Janssen & Shadlen, 2005; Trillenberg et al., 2000). Interval-based temporal predictions can align neural oscillations to sensory inputs by the phase-resetting oscillations whose period matches the interval (Haegens & Zion Golumbic, 2018; Rimmele et al., 2018; Schroeder & Lakatos, 2009). The phase of endogenous oscillations can be modulated by the internal representation of the temporal statistics of events (Breska & Deouell, 2014; Daume et al., 2021; Herbst & Obleser, 2019) and is sensitive to the shape of the implicit foreperiod distribution (Cravo et al., 2011).

The distinction between rhythmic and non-rhythmic temporal predictions has important consequences for the postulated neural mechanisms. While rhythmic temporal motifs are typically salient enough to be extracted in a bottom-up manner, interval-based predictions range on a continuum from strictly implicit, bottom-up modulations of temporal statistics (Bueti et al., 2010; Cravo et al., 2011; Herbst & Obleser, 2017) to cued top-down temporal orientation (Nobre & van Ede, 2017). Whether oscillatory responses observed during rhythmic stimulation truly reflect an endogenous representation of temporal predictions, or reflect the exogenous structure of the inputs is thus an open question.

The difficulties encountered in Section 2 highlight the need for empirical explorations that tease out the origins of temporal predictions, as well as their relation to endogenous attention and sensory saliency. These issues are a prelude to the next section, in which we describe the tension between endogenous and exogenous rhythms. The first issue that arises is the very existence of spontaneous neural oscillations (i.e., in the absence of driving external stimuli), which stems from skepticism about the functional relevance of brain rhythms to cognition. A second issue results from the need to distinguish the elicitation of neural oscillation from the modulation of ongoing oscillations (as posited and questioned by neural entrainment).

3 The tension in synchronization

While the previous section outlined a complex system evolved to align predictively with incoming rhythms, here we show that the perceiver may resist such alignment according to its own internal processing. Our brains exert endogenous control over how it perceives: deciding which beats are most important, imagining other melodies that might go well with a piece, or perceiving different meters where multiple interpretations are possible. At the same time, we must also consider that stimuli are not isolated; they are generated intentionally by producers, with the aim of eliciting mental constructs in the listeners (e.g., emotions or semantic concepts) through manipulations of internal processes. The dynamism in the musical interaction lies not in fully successful synchronization, but in the attempt to control and align to a constantly shifting experience.

3.1 Modulation of endogenous oscillations: uncoupled coupling

If brain rhythms remain under individual's endogenous and volitional control, a straightforward prediction is that one should be able to elicit a neural oscillation at will. This was verified in the elegant EEG study of Nozaradan and colleagues (Nozaradan et al., 2011): human participants listening to rhythmic auditory stimuli (2.4 Hz periodicity) were asked to imagine a binary meter (1.2 Hz), or a ternary meter (0.8 Hz), aligned to the 2.4 Hz beat, but to refrain from overt motor movements or counting. As predicted by entrainment, EEG activity showed clear spectral peaks at the beat frequency corresponding to the acoustic inputs. More interestingly, it also showed spectral peaks at the frequencies of the mentally generated meter. These findings are essential in showing that neural entrainment is not solely a bottom-up brain response to exogenous periodicities, it also reflects an endogenous temporal structure that is experienced by the listener. In fact, oscillatory dynamics can be measured in the absence of exogenous stimulation and resonate during and after rhythmic (Cason et al., 2015; Falk et al., 2017; Pesnot Lerousseau et al., 2021) and non-rhythmic (Teng et al., 2018) stimulation. Such intrinsic spontaneous oscillators support the biological and psychological needs for adaptability and stability.

Furthermore, endogenous control of brain responses can regulate sensory entrainment in subtle ways that are essential to psychological timing. Neural oscillations are not pure sinusoids and rather display non-stationarities, like bursting changes or slow changes in phase response. An example of non-stationarity in the oscillatory phase response has been argued to maintain an endogenous internal metric for the representation of conscious timing (Kösem et al., 2014). In this MEG study, desynchronized audiovisual stimuli were rhythmically presented to elicit a peak following response at the entrainment rate of 1 Hz (δ). The presentation of desynchronized audiovisual stimuli yielded a temporal recalibration phenomenon, in which participants' perception of audiovisual simultaneity shifted. As MEG records the mean response of latent neural oscillators, the entrained 1 Hz response should remain constantly in-phase and temporally aligned with the rhythmic presentation of the stimuli (Bauer et al., 2020; Thut et al., 2011). Contrary to this prediction, the phase of the auditory response systematically changed in a manner that linearly predicted participants' time perception. These findings suggested that

while the brain may stabilize its internal time metric through external entrainment, the need to maintain a unitary perception of the most probable multisensory event may endogenously adjust the phase response. Consistent with this observation, the (partial) independence of endogenous temporal references with exogenous temporal statistics suggest that some properties of entrained oscillations are physiologically independent from the external inputs, and exist beyond the passive tracking of sensory signals (Obleser & Kayser, 2019).

Altogether, these results support the idea that neural oscillations play a mechanistic role in information processing and that the phase of neural oscillations plays a tangible role in timing under endogenous control. In this sense, what fundamentally distinguishes animals from other living matter is their capacity to generate timing in manner that remains coherent with, but partially independent from, the temporal properties of their environment (van Wassenhove, 2016, 2017).

3.2 Synchronizing with others: Rhythms, almost

Given the complex predictive apparatus, why can we not perfectly predict rhythmic behaviors of others? Surprising stimuli attract exogenous attention (Theeuwes, 1991): unexpected stimuli, like a sudden thunder or a bright flash, automatically grab our attention in a bottom-up manner. The same occurs in less dramatic scenarios: surprising words in a sentence ("I like my coffee with cream and socks") attract more attention than predictable ones (Itti & Baldi, 2009; Zarcone et al., 2016). The subtle interplay between anticipation and surprisal is essential in music (Huron, 2008). Surprisal, formulated as the error between what is predicted and what occurs, is also a mathematical definition of information content. Inputs that differ from our predictions contain new information (Shannon, 1948).

A common example of the utility of surprisal in music is when performers in dance or music vary the timing of their expression. Unscripted and expressive shifts in musical timing, commonly referred to as rubato, communicate to the audience added information that they should prepare to clap, or to respond with their own melody depending on the environment. On the other hand, a long pause in the middle of a piece may generate surprise or tension in the listener, attracting their attention to the next phrase. A good performer in dance or music varies the timing of their performance. These unscripted and expressive shifts in musical timing, often referred to as rubato (Clarke, 1989; Todd, 1985) further engage the listener and often evoked emotional experiences. These behaviors, therefore, differ from the ubiquitous biological rhythms in our environment because they explicitly manipulate the perceiver, rendering them maximally receptive to new information.

How this behavior should relate to neural oscillations remains a matter of debate. Recent proposals suggest that such synchronization can align with imprecise expressive rhythms (Doelling & Assaneo, 2021). In this sense, it may be in the producer's interest to maintain enough of a rhythm that synchronization is meaningful, while also introducing enough complexity to avoid stationarity. Furthermore, a single oscillator is likely too simple a mechanism to support the full complexity of rhythms that we experience in dance and music. Here is where the motor cortex, with its ability to flexibly simulate temporal trajectories at a wide array of time scales, may come into play. The interplay between motor trajectories and oscillatory synchronization may form the basis of this interaction between producer and listener.

From a temporal prediction perspective, a tension arises between the listener (audience) and the producer (expresser): the former thrives on predicting the timing of events; the latter thrives on attention-grabbing surprise. The result is a feedback-loop system, which must

constantly progress: listeners improve their internal model of the producers, the improvement in (non-)rhythmic predictions diminishes surprisal, which forces the invention of new surprising techniques. Despite – or perhaps because of – our endogenous biases towards predicting rhythms and patterns, humans tend to avoid perfect predictability. In music, listeners experience the desire to move, and greater pleasure, when the rhythmic complexity is at a medium level (Matthews et al., 2019; Witek et al., 2014): too complex, and the stimulus becomes unpredictable, difficult to track—too simple, and it becomes overly predictable. While it may appear that the "goal" of neural synchronization is to become fully aligned with the outside world, human behavior suggests a different approach. The tension between our internal rhythms and those of our environment is an important feature of our perception, helping us to learn new rhythms and/or to contribute our own creativity to the process.

Conclusions

In this chapter, we discussed the interplay of rhythms rooted in neurobiology and emerging in human expression as tension between endogenous brain rhythms and exogenous entrainment by sensorimotor inputs. Tension is reminiscent of homeostasis and the need to maintain an equilibrium in face of ever-changing environments. For this equilibrium to remain steady, biological rhythms under automatic and voluntary control require efficient and reliable circuits. In the absence of entrainment to external clocks, intrinsic rhythms maintain their course, displaying distinct, endogenously defined periodicities, including neuronal ones (Webb et al., 2009). It is noteworthy that some neuronal circuits, such as central pattern generators, are specialized in producing rhythmic behaviors like walking, swimming, breathing, or chewing *in the absence of rhythmic inputs* (Marder & Bucher, 2001). Life is also unquestionably a matter of rhythms and most bodily rhythms exist independently of externally imposed stimulations: breathing or beating hearts consist in the periodic alternations of physiological inflows and outflows. In many species, synchronization and entrainment are not solely defined by passive universal time constants, they are also driven by social interactions (Aschoff et al., 1971; Bloch, 2010; Eban-Rothschild & Bloch, 2012). While this may be unsurprising for dancers, musicians and cognitive neuroscientists, that social synchronization may affect basic biological functioning such as clocking mechanisms (even in passive biological clocks) is intriguing and essential. This is the seed for collective behaviors whose complexity can be increased by some of the fundamental mechanisms we have described – active sensing, dynamic attending, entrainment, temporal predictions, and ultimately, synchronization. All lead to the creativity of human productions.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure & Function*, *218*(3), 611–643. https://doi.org/10.1007/s00429-012-0475-5

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, *25*(9), 3077–3085.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Aschoff, J., Fatranska, M., Giedke, H., Doerr, P., Stamm, D., & Wisser, H. (1971). Human circadian rhythms in continuous darkness: Entrainment by social cues. *Science*, *171*(3967), 213–215.

Bar, M. (2011). *Predictions in the brain: Using our past to generate a future*. Oxford University Press.

Bareš, M., Apps, R., Avanzino, L., Breska, A., D'Angelo, E., Filip, P., Gerwig, M., Ivry, R. B., Lawrenson, C. L., & Louis, E. D. (2019). Consensus paper: Decoding the contributions of

the cerebellum as a time machine. From neurons to clinical applications. *The Cerebellum*, *18*(2), 266–286.

Bauer, A.-K. R., Debener, S., & Nobre, A. C. (2020). Synchronisation of Neural Oscillations and Cross-modal Influences. *Trends in Cognitive Sciences*, *24*(6), 481–495. https://doi.org/10.1016/j.tics.2020.03.003

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., Emerson, R. G., & Schroeder, C. E. (2011). Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *Journal of Neuroscience*, *31*(9), 3176–3185. https://doi.org/10.1523/JNEUROSCI.4518-10.2011

Bloch, G. (2010). The Social Clock of the Honeybee. *Journal of Biological Rhythms*, *25*(5), 307–317. https://doi.org/10.1177/0748730410380149

Bouwer, F. L., Honing, H., & Slagter, H. A. (2020). Beat-based and memory-based temporal expectations in rhythm: Similar perceptual effects, different underlying mechanisms. *Journal of Cognitive Neuroscience*, *32*(7), 1221–1241.

Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *Journal of Cognitive Neuroscience*, *26*(7), 1555–1571. https://doi.org/10.1162/jocn_a_00564

Brown, H., Friston, K. J., & Bestmann, S. (2011). Active inference, attention, and motor preparation. *Frontiers in Psychology*, *2*, 218.

Bueti, D., Bahrami, B., Walsh, V., & Rees, G. (2010). Encoding of Temporal Probabilities in the Human Brain. *Journal of Neuroscience*, *30*(12), 4343–4352. https://doi.org/10.1523/JNEUROSCI.2254-09.2010

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, *107*(37), 16048–16053.

Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.

Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling Brain Size, Keeping Timing: Evolutionary Preservation of Brain Rhythms. *Neuron*, *80*(3), 751–764. https://doi.org/10.1016/j.neuron.2013.10.002

Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta Psychologica*, *155*, 43–50. https://doi.org/10.1016/j.actpsy.2014.12.002

Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, *51*(1), 2–9.

Cole, S. R., & Voytek, B. (2017). Brain oscillations and the importance of waveform shape. *Trends in Cognitive Sciences*, *21*(2), 137–149.

Cole, S. R., & Voytek, B. (2019). Cycle-by-cycle analysis of neural oscillations. *Journal of Neurophysiology*, *122*(2), 849–861. https://doi.org/10.1152/jn.00273.2019

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2011). Endogenous modulation of low frequency oscillations by temporal expectations. *Journal of Neurophysiology*, *106*(6), 2964–2972.

Daume, J., Wang, P., Maye, A., Zhang, D., & Engel, A. K. (2021). Non-rhythmic temporal prediction involves phase resets of low-frequency delta oscillations. *NeuroImage*, *224*, 117376. https://doi.org/10.1016/j.neuroimage.2020.117376

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Doelling, K. B., & Assaneo, M. F. (2021). Neural oscillations are a start toward understanding brain activity rather than the end. *PLoS Biology*, *19*(5), e3001234.

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences*, *116*(20), 10113–10121.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233–E6242.

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience*, *25*(36), 8259–8266.

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Dupré la Tour, T., Tallot, L., Grabot, L., Doyère, V., van Wassenhove, V., Grenier, Y., & Gramfort, A. (2017). Non-linear auto-regressive models for cross-frequency coupling in neural time series. *PLOS Computational Biology*, *13*(12), e1005893. https://doi.org/10.1371/journal.pcbi.1005893

Eban-Rothschild, A., & Bloch, G. (2012). Social influences on circadian rhythms and sleep in insects. *Advances in Genetics*, *77*, 1–32.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat Rev Neurosci*, *2*(10), 704–716. https://doi.org/10.1038/35094565

Falk, S., Lanzilotti, C., & Schön, D. (2017). Tuning Neural Phase Entrainment to Speech. *Journal of Cognitive Neuroscience*, *29*(8), 1378–1389. https://doi.org/10.1162/jocn_a_01136

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. *Current Biology*, *23*(24), 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019–1021.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and Gamma Rhythms in Human Auditory Cortex during Musical Beat Processing. *Annals of the New York Academy of Sciences*, *1169*(1), 89–92. https://doi.org/10.1111/j.1749-6632.2009.04779.x

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *Journal of Neuroscience*, *32*(5), 1791–1802. https://doi.org/10.1523/JNEUROSCI.4107-11.2012

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511.

Godøy, R. I., Song, M., Nymoen, K., Haugen, M. R., & Jensenius, A. R. (2016). Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research*. https://www.tandfonline.com/doi/full/10.1080/09298215.2016.1184689

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hausdorff, J. M., Purdon, P. L., Peng, C. K., Ladin, Z., Wei, J. Y., & Goldberger, A. L. (1996). Fractal dynamics of human gait: Stability of long-range correlations in stride interval fluctuations. *Journal of Applied Physiology*, *80*(5), 1448–1457. https://doi.org/10.1152/jappl.1996.80.5.1448

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Herbst, S. K., Fiedler, L., & Obleser, J. (2018). Tracking Temporal Hazard in the Human Electroencephalogram Using a Forward Encoding Model. *Eneuro*, *5*(2), ENEURO.0017-18.2018. https://doi.org/10.1523/ENEURO.0017-18.2018

Herbst, S. K., & Obleser, J. (2017). Implicit variations of temporal predictability: Shaping the neural oscillatory and behavioural response. *Neuropsychologia*, *101*, 141–152. https://doi.org/10.1016/j.neuropsychologia.2017.05.019

Herbst, S. K., & Obleser, J. (2019). Implicit temporal predictability enhances pitch discrimination sensitivity and biases the phase of delta oscillations in auditory cortex. *NeuroImage*, *203*, 116198. https://doi.org/10.1016/j.neuroimage.2019.116198

Herrmann, B., Henry, M. J., Grigutsch, M., & Obleser, J. (2013). Oscillatory phase dynamics in neural entrainment underpin illusory percepts of time. *Journal of Neuroscience*, *33*(40), 15799–15809.

Ho, H. T., Leung, J., Burr, D. C., Alais, D., & Morrone, M. C. (2017). Auditory sensitivity and decision criteria oscillate at different frequencies separately for the two ears. *Current Biology*, *27*(23), 3643–3649.

Huron, D. (2008). *Sweet anticipation: Music and the psychology of expectation*. MIT press.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Jacoby, N., & McDermott, J. H. (2017). Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Current Biology*, *27*(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience*, *8*(2), 234–241. https://doi.org/10.1038/nn1386

Jasper, H. H. (1948). Charting the Sea of Brain Waves. *Science*, *108*(2805), 343–347. JSTOR.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323–355. https://doi.org/10.1037/0033-295X.83.5.323

Kononowicz, T. W., Roger, C., & van Wassenhove, V. (2019). Temporal metacognition as the decoding of self-generated brain dynamics. *Cerebral Cortex*, *29*(10), 4366–4380.

Kösem, A., Gramfort, A., & van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *Neuroimage*, *92*, 274–284.

Kösem, A., & Van Wassenhove, V. (2012). Temporal structure in audiovisual sensory selection. *PLoS One*, *7*(7), e40936.

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology*, *29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004.

Large, E., & Jones, M. (1999). The Dynamics of Attending: How People Track Time-Varying Events. *Psychological Review*, *106*(1), 119–159.

Llinás, R. R. (1988). The intrinsic electrophysiological properties of mammalian neurons: Insights into central nervous system function. *Science*, *242*(4886), 1654–1664.

Local, J. K., Kelly, J., & Wells, W. H. (1986). Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics*, 411–437.

London, J. (2012). *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford University Press.

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Marder, E., & Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Current Biology*, *11*(23), R986–R996. https://doi.org/10.1016/S0960-9822(01)00581-4

Massion, J. (1994). Postural control system. *Current Opinion in Neurobiology*, *4*(6), 877–887. https://doi.org/10.1016/0959-4388(94)90137-6

Matthews, T. E., Witek, M. A. G., Heggli, O. A., Penhune, V. B., & Vuust, P. (2019). The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PLOS ONE*, *14*(1), e0204539. https://doi.org/10.1371/journal.pone.0204539

Miller, J. E., Carlson, L. A., & McAuley, J. D. (2013). When what you hear influences when you see: Listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science*, *24*(1), 11–18.

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *Journal of Neuroscience*, *36*(8), 2342–2347.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162. https://doi.org/10.1037/0033-2909.89.1.133

Nobre, A. C., & van Ede, F. (2017). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, *19*. https://ora.ox.ac.uk/objects/uuid:2c60d934-4ab2-4d87-8346-3a2fcd6e3a05

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *Journal of Neuroscience*, *31*(28), 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *Journal of Neuroscience*, *32*(49), 17572–17581. https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004

Panzeri, S., Brunel, N., Logothetis, N. K., & Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends in Neurosciences*, *33*(3), 111–120. https://doi.org/10.1016/j.tins.2009.12.001

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.

Pesnot Lerousseau, J., Trébuchon, A., Morillon, B., & Schön, D. (2021). Frequency selectivity of persistent cortical oscillatory responses to auditory rhythmic stimulation. *BioRxiv*, 834226. https://doi.org/10.1101/834226

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic Prototypes Across CulturesA Comparative Study of Tapping Synchronization. *Music Perception*, *36*(1), 1–23. https://doi.org/10.1525/mp.2018.36.1.1

Pöppel, E. (1971). Oscillations as possible basis for time perception. *Stud Gen (Berl)*, *24*(1), 85–107.

Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, *24*(1), 13–24.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882.

Ross, A. (2007). *The rest is noise: Listening to the twentieth century*. Macmillan.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176. https://doi.org/10.1016/j.conb.2010.02.010

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience*, *30*(41), 13578–13585.

Stokes, R. C., Venezia, J. H., & Hickok, G. (2019). The motor system's [modest] contribution to speech perception. *Psychonomic Bulletin & Review*, *26*(4), 1354–1366. https://doi.org/10.3758/s13423-019-01580-2

Su, Y.-H. (2016). Visual tuning and metrical perception of realistic point-light dance movements. *Scientific Reports*, *6*(1), 22774. https://doi.org/10.1038/srep22774

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409.

Teng, X., Tian, X., Doelling, K., & Poeppel, D. (2018). Theta band oscillations reflect more than entrainment: Behavioral and neural evidence demonstrates an active chunking process. *European Journal of Neuroscience*, *48*(8), 2770–2782. https://doi.org/10.1111/ejn.13742

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. https://doi.org/10.3758/BF03211619

Thut, G., Schyns, P., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology*, *2*, 170.

Todd, N. (1985). A Model of Expressive Timing in Tonal Music. *Music Perception*, *3*(1), 33–57. https://doi.org/10.2307/40285321

Todorovic, A., Schoffelen, J.-M., Van Ede, F., Maris, E., & De Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PLoS One*, *10*(3), e0120288.

Trillenberg, P., Verleger, R., Wascher, E., Wauschkuhn, B., & Wessel, K. (2000). CNV and temporal uncertainty with 'ageing' and 'non-ageing' S1–S2 intervals. *Clinical Neurophysiology*, *111*(7), 1216–1226. https://doi.org/10.1016/S1388-2457(00)00274-1

Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053.

Varela, F. J. (1999). The specious present: A neurophenomenology of time consciousness. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, *64*, 266–329.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Rev. Neurosci.*, *2*, 229–239. https://doi.org/10.1038/35067550

Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, *90*(2), 858–865. https://doi.org/10.1121/1.401953

Wang, X.-J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, *90*(3), 1195–1268. https://doi.org/10.1152/physrev.00035.2008

Webb, A. B., Angelo, N., Huettner, J. E., & Herzog, E. D. (2009). Intrinsic, nondeterministic circadian rhythm generation in identified mammalian neurons. *Proceedings of the National Academy of Sciences*, *106*(38), 16493–16498. https://doi.org/10.1073/pnas.0902768106

Whittingstall, K., & Logothetis, N. K. (2009). Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron*, *64*(2), 281–289.

Wilbur, R. B., & Nolkn, S. B. (1986). The Duration of Syllables in American Sign Language. *Language and Speech*, *29*(3), 263–280. https://doi.org/10.1177/002383098602900306

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, *12*(1), 1–24.

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, Body-Movement and Pleasure in Groove Music. *PLOS ONE*, *9*(4), e94446. https://doi.org/10.1371/journal.pone.0094446

Wöllner, C., & Cañal-Bruland, R. (2010). Keeping an eye on the violinist: Motor experts show superior timing consistency in a visual perception task. *Psychological Research*, *74*(6), 579–585. https://doi.org/10.1007/s00426-010-0280-9

Woodrow, H. (1914). The measurement of attention. *The Psychological Monographs*, *17*(5), i–158. https://doi.org/10.1037/h0093087

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences*, *109*(9), 3593–3598.

Zarcone, A., van Schijndel, M., Vogels, J., & Demberg, V. (2016). Salience and Attention in Surprisal-Based Accounts of Language Processing. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00844