Neural networks for beat perception in musical rhythm

Edward Large [1,*]**, Jorge Herrera** [2] **and Marc Velasco** [3]

[1]*Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA*
[2]*Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, CA, USA*
[3]*Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA*

Correspondence*:
Edward Large
Department of Psychological Sciences, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, 06269-1020, CT, USA, edward.large@uconn.edu

ABSTRACT

Entrainment of cortical rhythms to acoustic rhythms has been hypothesized to be the neural correlate of pulse and meter perception in music. Dynamic attending theory first proposed synchronization of endogenous perceptual rhythms nearly forty years ago, but only recently has the pivotal role of neural synchrony been demonstrated. Significant progress has since been made in understanding the role of neural oscillations and the neural structures that support synchronized responses to musical rhythm. Synchronized neural activity has been observed in auditory and motor networks, and has been linked with attentional allocation and movement coordination. Here we describe a neurodynamic model that shows how self-organization of oscillations in interacting sensory and motor networks could be responsible for the formation of the pulse percept in complex rhythms. In a pulse synchronization study, we test the model's key prediction that pulse can be perceived at a frequency for which no spectral energy is present in the amplitude envelope of the acoustic rhythm. The result shows that participants perceive the pulse at the theoretically predicted frequency. This model is one of the few consistent with neurophysiological evidence on the role of neural oscillation, and it explains a phenomenon that other computational models fail to explain. Because it is based on a canonical model the predictions hold for an entire family of dynamical systems, not only a specific one. Thus, this model provides a theoretical link between oscillatory neurodynamics and the induction of pulse and meter in musical rhythm.

Keywords: beat perception, neural resonance, musical rhythm, dynamical systems, neural networks

1 PERCEPTION OF PULSE AND METER IN MUSICAL RHYTHMS

The sounds that humans use for communication are temporally structured sequences of events, such as musical notes or speech syllables. Rhythm refers to the pattern of timing and stress in the amplitude envelope of an acoustic sequence. Musical rhythms are usually perceived to have a pulse, or basic beat, in the approximate frequency range $0.5 - 4$ Hz (London, 2004). Meter is a perceived temporal structure that includes the pulse frequency, as well as slower beat frequencies that accent pulse cycles (< 2 Hz),

27 and higher beat frequencies that subdivide the pulse (4 – 8 Hz) (London, 2004; Lerdahl and Jackendoff,
28 1983). Neural resonance theory (Large and Snyder, 2009) hypothesizes that pulse and meter correspond to
29 neural rhythms that synchronize with acoustic rhythms, influencing temporal expectancy, attention, and
30 movement coordination. Theoretical approaches to understanding perceived structure have been based on
31 neurodynamic models of neural oscillation (for a review, see Large, 2008). However, only recently has a
32 clear picture begun to emerge regarding synchronization of neural oscillations and the neural structures
33 that support responses to musical rhythm.

34   In this paper, we summarize current knowledge about the synchronization of neural rhythms to musical
35 rhythms and outline a neurodynamic model of pulse perception based on entrainment of neural oscillation.
36 First, in §2, we present a brief overview of the main theories and experimental findings related to musical
37 pulse and meter. We discuss the potential function of neural oscillations in establishing the perceived
38 temporal structure of complex musical rhythms. In §3, we sketch a neurodynamic model of pulse perception
39 based on the interaction between oscillatory neural networks. The model incorporates the basic findings
40 of the past 20 or so years and makes a key prediction about the formation of the pulse percept. In §4, we
41 evaluate the fundamental prediction of the theory, that perceived temporal structures may correspond to
42 frequencies that are not physically present in the amplitude envelope. Finally, in §5, we discuss remaining
43 important questions regarding the link between pulse perception and neural resonance.

2   NEURAL RESONANCE TO MUSICAL RHYTHMS

44 Neural population rhythms, as observed in the local field potential (LFP), the electroencephalogram (EEG),
45 and the magnetoencephalogram (MEG), are cyclical fluctuations of baseline neuronal activity that can
46 be observed in neocortical and thalamic regions of the brain (Steriade et al., 1993; Buzsáki and Draguhn,
47 2004; Slézia et al., 2011). Thalamocortical oscillations exhibit $1/f$ frequency spectra with peaks in specific
48 frequency bands, including delta (~1 – 4 Hz), theta (~4 – 8 Hz), beta (~13 – 30 Hz), and gamma (~30 – 70
49 Hz) (Buzsáki, 2006). The frequency range of musical pulse (London, 2004) corresponds nicely with the
50 delta band, while faster metrical frequencies fall within the theta band, and slower metrical frequencies
51 occupy the sub-delta range (Musacchia et al., 2013, see Figure 1). Jones originally hypothesized that
52 endogenous perceptual rhythms synchronize with temporally structured sequences, generating expectancies
53 for future events (Jones, 1976). Dynamic attending theory (DAT) (Jones, 1976; Jones and Boltz, 1989;
54 Large and Jones, 1999) addressed the issue of how neural rhythms may be exploited by an organism
55 to enable attentional coordination with the dynamic external world (for a review, see Jones, 2008). The
56 sensory-motor theory of rhythm perception (Todd, 1994, 1995; **?**), hypothesized that rhythm and pulse
57 perception involve a sensory representation of the input as well as a motor representation of the body. A
58 related theoretical framework that accounts for the predictive role of the motor system has recently been
59 proposed (Schubotz, 2007). Converging lines of investigation have provided evidence supporting both
60 theoretical perspectives, and suggest that interaction between sensory and motor regions of the brain may
61 provide a mechanism for predicting sequence timing (Schroeder et al., 2010). In this paper, we further
62 develop a neural resonance theory of rhythm perception, hypothesizing that oscillatory interactions between
63 sensory and motor areas are sufficient to give rise to percepts of pulse and meter in complex musical
64 rhythms (Large and Snyder, 2009).

65   Behavioral evidence for dynamic attending in adults and infants comes from time discrimination and
66 pitch discrimination studies that revealed an advantage for auditory sequences that induce a metric percept,
67 using both sensitivity and reaction time measures (Large and Jones, 1999; McAuley and Kidd, 1995;
68 Yee et al., 1994; Jones and Yee, 1997; Barnes and Jones, 2000; Jones and McAuley, 2005; Jones et al.,

69 2006; Bergeson and Trehub, 2006). Such findings are consistent with early-developing predispositions for
70 temporal regularity, and a preference for more stable temporal organizations, as predicted by nonlinear
71 resonance (Large, 2008). In support of sensory-motor theory, a fundamental sound-movement interaction
72 was demonstrated in infants and adults such that vestibular stimulation can influence whether an ambiguous
73 pattern is perceived in duple or triple meter (Phillips-Silver and Trainor, 2005, 2007). The top-down
74 influence of motor networks on sensory representations has been recently studied, showing how the former
75 can "sharpen" the temporal selection of auditory information (Morillon et al., 2014, 2015). In addition, the
76 finding of perceptual narrowing in infants demonstrated that meter perception is plastic and depends upon
77 musical enculturation (Hannon and Trehub, 2005).

78   EEG and MEG studies have directly tested the role of endogenous oscillation. One EEG study revealed
79 that fluctuations in induced beta- and gamma-band power synchronized with periodic and metrical
80 rhythms, revealing both sensory-driven and anticipatory responses to tones (Snyder and Large, 2005,
81 Figure 2a). A related MEG study also found sensory and anticipatory responses in the beta- and gamma-
82 band (Fujioka et al., 2009). When subjects were instructed to impose a subjective meter on a periodic
83 stimulus, subharmonic responses were observed in induced beta band MEG activity, closely resembling
84 those produced by physical accents (Iversen et al., 2009). One EEG study reported synchronization of
85 delta rhythms, phasic responses in theta, and augmented phase synchronization throughout the beta/gamma
86 range, modulated by the stimulus periodicity (Will and Berg, 2007). Another study revealed entrainment of
87 the delta rhythm to the onset of target tones, and reaction times that correlated with the phase of the delta
88 band oscillation at target onset, directly supporting the role of neural oscillation in attending (Stefanics
89 et al., 2010). The steady-state evoked potential (SS-EP) technique revealed that a periodic rhythm elicited
90 a sustained response in the delta band, and meter imagery elicited an additional subharmonic resonance
91 corresponding to the metric interpretation (Nozaradan et al., 2011, Figure 2b). Another study showed
92 that complex rhythms elicited multiple SS-EPs in the EEG spectrum at frequencies corresponding to the
93 rhythmic pattern envelope, and the amplitude of the SS-EPs at pulse and meter frequencies was selectively
94 enhanced, suggesting a role for neural oscillations in pulse and meter induction (Nozaradan et al., 2012).
95 Thus, delta, beta and gamma band responses to auditory rhythms observed in both in EEG and MEG in
96 humans have confirmed predictions of dynamic attending theory. Given the role of beta in motor processing
97 and long-range intra-cortical interaction, these findings are consistent with the idea that the motor system
98 influences the perception of sound, even in the absence of overt movement.

99   Functional magnetic resonance imaging (fMRI) studies of musical rhythm perception have clearly
100 demonstrated that listening to musical rhythms recruits both auditory and motor areas of the brain (Chen
101 et al., 2008a, Figure 2c). A functionally connected network is implicated in extracting higher-order features
102 of a rhythm's temporal structure, with the dorsal premotor cortex mediating auditory-motor interactions,
103 and functional coupling of auditory-motor networks observed even in perception tasks without a motor
104 component (Chen et al., 2008b). Such studies have implicated the basal ganglia (Grahn and Rowe, 2009),
105 cerebellum and supplementary motor area (Chen et al., 2008a; Bengtsson et al., 2009), as well as the dorsal
106 premotor cortex and right frontal lobe (Bengtsson et al., 2009). Functional connectivity analysis implicates a
107 cortico-subcortical network including the putamen, SMA, and PMC for the analysis of temporal sequences,
108 especially under conditions that may require internal generation of the pulse (Grahn and Rowe, 2009,
109 Figure 2d). Basal ganglia may play an important functional role in the formation of the pulse percept
110 (Chapin et al., 2010; Kung et al., 2013) or pulse prediction (Grahn and Rowe, 2013). Although functional
111 imaging cannot resolve rhythmic time scales, MEG source analysis has revealed beta band interactions in
112 auditory and motor networks during musical rhythm processing, implying oscillatory interactions among
113 auditory and motor cortices as well as the cerebellum, thalamus, and parahippocampal gyrus (Fujioka et al.,

114 2012, Figure 2e). Synchronization of beta suppression with the auditory input (Fujioka et al., 2012) and
115 the spatial overlap of the beta network with the striato-thalamo-cortical network implicated by functional
116 imaging suggests that the mechanism of pulse and meter perception is fundamentally oscillatory. Moreover,
117 in macaque monkeys, beta- and gamma-band oscillations measured using LFPs during synchronization-
118 continuation tasks suggest differential roles in rhythmic tapping versus stimulus processing (Bartolo et al.,
119 2014). Observations of beta oscillations in basal ganglia during synchronization and continuation further
120 support their role in the striato-thalamo-cortical circuit during control of rhythmic behavior—for a recent
121 review, see Merchant et al. 2015a—(Bartolo and Merchant, 2015). Interaction between sensory and motor
122 regions of the brain has been found not only when listening to music, but also in speech perception and
123 production (Rauschecker and Scott, 2009).

124   A summary of this scenario is outlined in Figure 3. Interaction of excitatory and inhibitory neuronal
125 populations gives rise to population rhythms throughout the brain (Brunel, 2003; Börgers and Kopell,
126 2003; Buzsáki and Draguhn, 2004; Stefanescu and Jirsa, 2008), including sensory and motor networks.
127 When sensory stimuli are presented in a periodic pattern, ambient delta-band oscillations entrain to the
128 structure of the stimulus stream (Will and Berg, 2007; Nozaradan et al., 2011, 2012). Fluctuations in
129 beta- and gamma-band rhythms synchronize as well, consistent with an oscillatory hierarchy in auditory
130 cortex (Lakatos et al., 2005). Neuronal entrainment emerges rapidly and facilitates behavioral responses
131 (Stefanics et al., 2010). The perception of pulse and meter involves broadly distributed motor systems
132 (Chen et al., 2008a; Grahn and Rowe, 2009; Kung et al., 2013; Grahn and Rowe, 2013) which respond with
133 synchronized fluctuations in gamma- and beta-band amplitude (Fujioka et al., 2012), enabling coordination
134 of perception and rhythmic movements with musical rhythms (Chen et al., 2008b; Nozaradan et al., 2013).
135 Auditory-motor coupling is reciprocal (Phillips-Silver and Trainor, 2005, 2007), and connections within
136 and between sensory and motor systems are assumed to be plastic (Hannon and Trehub, 2005).

3 OSCILLATORY NETWORK INTERACTIONS COULD GIVE RISE TO MUSICAL PULSE PERCEPTION

137 A fundamental challenge to understanding a complex neural system such as the one just described lies in
138 the integration of data from multiple modalities and multiple levels of observation into a coherent systems
139 model of perception, attention and behavior. Musical rhythm provides a unique opportunity in this regard,
140 because acoustic signals are thought to drive entrainment of neuronal population oscillations. Nonlinear
141 oscillation in general, and neural oscillation in particular, have been widely studied in the mathematical
142 and physical sciences over the past 50 years, putting us in a unique position to link levels of observation
143 from spiking neurons to population dynamics and from population dynamics to perception and behavior,
144 because sophisticated mathematical tools are available for understanding the behavior of such systems.
145 Models of neuronal oscillations arising from the interaction of excitatory and inhibitory populations of
146 neurons can be used to study mechanisms of neural oscillation (Whittington et al., 2000), with implications
147 for many different aspects of neural information processing (Wilson and Cowan, 1973; Varela et al., 2001;
148 Hoppensteadt and Izhikevich, 2005; Koepsell et al., 2010; Ainsworth et al., 2012). Models may be driven
149 with external input (Brunel, 2003; Large et al., 2010), may include synaptic plasticity (Brunel, 2003;
150 Hoppensteadt and Izhikevich, 1996b), and predict a surprisingly rich repertoire of behaviors, including
151 steady states, Hopf bifurcations, double limit cycle bifurcations, bursting, and chaotic dynamics (Brunel,
152 2000; Stefanescu and Jirsa, 2008; Ledoux and Brunel, 2011).

153   Mathematical models can be useful in linking oscillation of high dimensional neuronal populations with
154 lower dimensional population-level models that capture much of the behavioral richness of observed in

high dimensional systems (Wilson and Cowan, 1973; Stefanescu and Jirsa, 2008), and are amenable to theoretical and computational analysis (Aronson et al., 1990; Hoppensteadt and Izhikevich, 1996a). We have proposed one such population-level model (Equation 1) (Large et al., 2010) to describe the dynamics of networks of neural oscillators with different natural frequencies responding to external stimulation:

$$\tau_i \frac{dx_i}{dt} = f_i(x_i, y_i, \lambda) + \epsilon p_i(x_1, y_1, \ldots, x_n, y_n, s(t), \lambda, \epsilon)$$
$$\tau_i \frac{dy_i}{dt} = g_i(x_i, y_i, \lambda) + \epsilon q_i(x_1, y_1, \ldots, x_n, y_n, s(t), \lambda, \epsilon)$$

(1)

Here, the variables x_i and y_i represent excitatory and inhibitory activities in the ith neuronal population respectively. The nonlinear functions f_i and g_i describe the intrinsic (uncoupled) dynamics of the excitatory and inhibitory subpopulations, and p_i and q_i describe interactions between neuronal subpopulations with external input $s(t)$. λ is a set of model parameters and τ_i is varied to create a frequency gradient as is found in many parts of the auditory system (Langner, 1992). ϵ is a small number that represents weak interaction (Large et al., 2010; Hoppensteadt and Izhikevich, 1996a).

This very general model implies a set of generic predictions about emergent neuronal oscillations under the influence of time-varying external input (e.g., musical rhythms) that holds under a broad set of assumptions (Large et al., 2010; Hoppensteadt and Izhikevich, 1996a). Equation 2 describes a canonical gradient frequency neural oscillator network derived from Equation 1 using normal form theory (Large et al., 2010), that has been used to model entrainment of perception, attention and behavior to rhythmic stimuli (Large, 2010).

$$\tau_i \frac{dz_i}{dt} = z_i(\alpha + \mathrm{i}2\pi + \beta_1 |z_i|^2 + \epsilon \beta_2 |z_i|^4 + \ldots) + \sum_{j \neq i}^{n} c_{ij} \mathcal{P}(\epsilon, z_j) \mathcal{A}(\epsilon, \bar{z}_i) + b_i s(t)$$

(2)

Here the roman i denotes the imaginary unit and z_i is the complex-valued state variable for the ith neural oscillator whose real and imaginary parts can be thought of as the activation of the excitatory and inhibitory subpopulations respectively. The ith oscillator's natural frequency is given by $f_i = 1/\tau_i$. The β's are nonlinear damping parameters, and the complete expansion of intrinsic terms describes a fully saturating nonlinearity (Large, 2010; Murdock, 2003). The parameter α controls the system's intrinsic behavior. $\alpha = 0$ is the critical point, above which the system exhibits spontaneous oscillation and below which damped oscillation, through an Andorov-Hopf bifurcation, illustrated in Figure 4a. The interaction of the external signal with the intrinsic oscillatory dynamics makes the key predictions in this model.

Nonlinear stimulus coupling predicts mode-locked responses of neural oscillators to the stimulus. Mode-locking is a generalization of phase-locking in which a periodic stimulus interacts with an intrinsic oscillatory dynamics of a neuron or neural circuit, causing k cycles of an oscillation to lock to m cycles of the stimulus, where k and m are integers, as shown in Figure 4b (for phase-locking, $k = m = 1$). Mode-locking predicts neural responses at harmonics, subharmonics, integer ratios, and combination frequencies of those present in a given rhythmic stimulus. Mode-locking in the canonical model is captured using a full expansion of resonant monomials (Large et al., 2010) which is expressed as a 'passive' nonlinear function $\mathcal{P}$ of input from another oscillator, z_j, multiplied by an 'active' nonlinear function $\mathcal{A}$ of the current state, z_i:

$$\mathcal{P}(\epsilon, z) = z + \sqrt{\epsilon}z^2 + \epsilon z^3 + \epsilon\sqrt{\epsilon}z^4 + \ldots = \frac{z}{1 - \sqrt{\epsilon}z} \tag{3}$$

$$\mathcal{A}(\epsilon, \bar{z}) = 1 + \sqrt{\epsilon}\bar{z} + \epsilon\bar{z}^2 + \epsilon\sqrt{\epsilon}\bar{z}^3 + \ldots = \frac{1}{1 - \sqrt{\epsilon}\bar{z}} \tag{4}$$

188 We use two gradient frequency networks to model the functional coupling of auditory-motor networks
189 observed in perception tasks without a motor component (Chen et al., 2008a; Velasco and Large, 2011).
190 The sensory network is intended to capture auditory cortical entrainment, while the motor network is
191 intended to capture the dynamics of a broadly distributed network including basal ganglia and cortical
192 areas. As illustrated in Figure 4c, the sensory network takes a rhythmic input, sends output to a motor
193 network, and the motor network send input back to the sensory network (for a detailed description, see
194 Velasco and Large, 2011). Connections within and between networks are assumed to be plastic and tuned
195 by musical enculturation (cf., Hannon and Trehub, 2005; Large, 2010; Velasco and Large, 2011).

196 Although the model makes only very general assumptions regarding underlying neural structures (e.g.,
197 Chen et al., 2008a), it makes strong commitments about the oscillatory dynamics of auditory-motor
198 interactions (Will and Berg, 2007; Fujioka et al., 2012; Nozaradan et al., 2013). The sensory oscillators
199 are tuned to operate near a Hopf bifurcation (Hoppensteadt and Izhikevich, 1997), as shown in Figure
200 4a; the motor oscillators are tuned to operate near a double limit cycle bifurcation (Velasco and Large,
201 2011; Izhikevich, 2000). The double limit cycle regime of the motor network means that the model can
202 capture synchronization-continuation behavior, continuing to produce rhythmic behavior after the stimulus
203 ceases (Wing and Kristofferson, 1973b,a). Here, however, we are interested in evaluating predictions
204 about synchronization (Velasco and Large, 2011). To predict mean field time series as observed in EEG
205 recordings (e.g., Will and Berg, 2007; Stefanics et al., 2010), we sum the output of all oscillators in each
206 network (Figure 5, left). To predict steady state evoked potentials (SS-EPs, e.g., Berens and Velasco, 2009)
207 we take a frequency analysis (DFT) of the mean field (Figure 5, right).

208 As shown in Figure 5a, for a periodic stimulus, both sensory and motor networks produce synchronized
209 oscillations at the pulse frequency, and generate harmonics (Repp, 2008) and subharmonics (Nozaradan
210 et al., 2011; Bolton, 1894; Vos, 1973). In the case of a complex rhythm, however, it becomes clear that
211 the two networks are doing something quite different from one another. The mean field time trace for the
212 sensory network represents the input rhythm rather faithfully, producing well-defined pulses at input event
213 times. By contrast, the motor network entrains at the pulse frequency. The rhythm itself contains no energy
214 at the pulse frequency (or its second subharmonic; DFT in Figure 5b, SS-EP, solid black), however, in the
215 motor network the strongest response is found at the pulse frequency. In other words, the development of
216 the pulse percept depends on the interaction of these two oscillatory systems.

217 This predicts that an oscillatory network interaction can lead to spontaneous pulse induction in complex
218 rhythms — even in the most extreme case of a rhythm for which there is no energy at the pulse frequency.
219 Thus, the theoretical prediction is that pulse may be perceived at a frequency that is not physically present
220 in the rhythmic stimulus (Large, 2010; Velasco and Large, 2011). The prediction for an isochronous rhythm
221 (Figure 5a) is consistent with observations of subharmonic resonance in EEG (Nozaradan et al., 2011),
222 however, in that EEG study subharmonic resonance was observed only when subjects were instructed to
223 imagine it. The behavior of the sensory network in response to the complex rhythm (Figure 5b) is consistent
224 with the observation of the enhancement of metrical frequencies in the SS-EP (Nozaradan et al., 2012).

225 However, the prediction of network interaction goes further in that the motor network produces metrical
226 frequencies that are completely absent from the rhythm. This is a strong prediction that has never been
227 tested.

228   This points up a critical issue in the interpretation of most of the empirical studies that have been
229 conducted to date. The problem that arises in the theoretical interpretation of synchronized periodic
230 movements or population-level neural recordings is that of linking observations of synchrony with intrinsic
231 dynamics of emergent population oscillations (Whittington et al., 2000). EEG oscillations, for example,
232 represent a remote consequence of periodic activity in the area of the brain between recording electrodes
233 and a reference electrode. To generate a signal distinct from background noise, this activity must be
234 manifest in a large enough proportion of neurons proximal to the electrode, implying a certain degree of
235 synchronization among neurons. However, there are at least two distinct mechanisms by which neuronal
236 synchrony can occur: it can be a consequence of an emergent population oscillation within the network
237 of cells under the electrode — with its own intrinsic dynamics — or it could be merely a consequence
238 of a common rhythmic input (Whittington et al., 2000). In neural resonance theory it is postulated that
239 both are present, i.e., a common rhythmic input entrains an emergent population oscillation. In this case,
240 interpretation of a synchronized neural responses is especially tricky, because a synchronized response
241 itself does not imply interaction of the input with an intrinsic oscillatory dynamics. However, demonstration
242 of a pulse percept at a frequency that is not physically present in a stimulus rhythm would directly implicate
243 intrinsic oscillatory dynamics in the perception of pulse.

4   DOES RHYTHMIC INPUT INTERACT WITH INTRINSIC DYNAMICS TO FORM THE PULSE PERCEPT?

244 Thus, the fundamental prediction to be tested is whether people will perceive a pulse frequency that
245 is not present in a rhythmic stimulus. This is a strong prediction, because if pulse perception could be
246 experimentally observed at a missing frequency, it would rule out the interpretation of synchronization due
247 merely to a common input frequency, and equivalently models of pulse induction based on linear resonance.
248 On the other hand, if a predicted missing pulse were not observed, it would falsify the hypothesis that pulse
249 induction occurs due interaction between a rhythmic stimulus and the intrinsic dynamics of endogenous
250 oscillatory networks.

251   Pulse is operationally defined as the frequency at which a listener coordinates a periodic movement
252 with a complex rhythm (Large, 2008). Thus, we asked participants to listen to eleven rhythms ranging
253 from isochronous to highly complex, where the most complex rhythms contained no spectral energy at
254 the pulse frequency predicted by the model, as shown in Figure 6. Participants were instructed to listen
255 to each rhythm until they heard a steady pulse, and then tap along with the rhythm at that rate. The
256 fourteen participants were primarily undergraduates from Florida Atlantic University, who received partial
257 credit in fulfillment of undergraduate psychology courses, one participant was a volunteer, and none were
258 professional musicians (8 males, 6 females; mean age 25.3; mean years of musical training: 3.2). All
259 participants provided informed consent, following guidelines set by the Institutional Review Board of
260 Florida Atlantic University.

261   Rhythms were presented at five levels of complexity (0 − 4) and at five different tempi (i.e., pulse
262 frequencies): 2.28 Hz (420 ms), 2.17 Hz (460 ms), 2 Hz (500 ms), 1.85 Hz (540 ms), and 1.72 Hz (580 ms).
263 The location of acoustic events on a metric grid grid was used to manipulate the level of rhythmic complexity
264 (i.e., syncopation, see Figure 6). Counting quarter-note beats as "strong" and eighth-note subdivisions as

265 "weak" (Lerdahl and Jackendoff, 1983), the level of syncopation was varied by manipulating the number of
266 events that fell on weak beats, as illustrated in Figure 6. Weak beats (subdivisions) can also be thought of
267 as existing in anti-phase relation to the in-phase strong beats. At complexity level 0, all eight events fell on
268 strong beats, at complexity level 1, one event fell on a weak beat and seven events fell on strong beats, and
269 so on. Level 4 was the greatest level of syncopation, with four events on weak beats and four events on
270 strong beats. Each pattern consisted of 8 sounded events (262 Hz sine waves, 50 ms duration, 5-ms rise and
271 fall times) laid out on an eight-beat metrical grid with binary subdivisions. Thus, the total duration of each
272 pattern was sixteen (eighth-note) beats, equivalent to eight quarter-note beats or two four-beat measures.
273 Each trial consisted of six pattern repetitions. Different combinations of tempo and rhythm were presented
274 in a pseudorandom order such that consecutive trials always had different rhythms, and different tempos.
275 Thus, participants were forced to find both the frequency and the phase of the pulse anew for each rhythm;
276 they could not simply tap at the same tempo throughout the experiment.

277 We measured instantaneous tapping frequency distributions to determine whether subjects induced a pulse
278 at the intended frequency. Instantaneous tapping frequency was computed as $1/\text{ITI}$ (ITI = inter-tap interval
279 in seconds) and tapping frequencies were normalized to a frequency of 2 Hz so they could be combined into
280 a single distribution at each level of complexity. Spectral analysis (DFT) of the stimulus rhythms (Figure
281 7a, black) shows that at the hypothetical pulse frequency amplitude decreases with increasing complexity.
282 At complexity level 4, the amplitude is precisely zero at 2 Hz and 1 Hz for each rhythm. Normalized
283 instantaneous tapping frequency (Figure 7a, red histogram) displays a main peak at the normalized pulse
284 frequency of 2 Hz for all rhythms at all levels of complexity, with lesser peaks at 1Hz and 0.5Hz, and
285 for some rhythms, a diffuse peak around 4 Hz. Thus, the participants most often tapped the predicted
286 pulse frequency even for the most complex rhythms, which had no spectral amplitude at that frequency.
287 The distributions for the higher levels of complexity (levels 3 and 4) were significantly different from the
288 distributions for the lower complexity levels (Kolmogorov-Smirnov, $D(N = 101) = 0.24, p < .01$). The
289 change appeared to be due to a decrease in the proportion of taps at 2 Hz and an increase in the proportion
290 of (unsynchronized) taps between 1 and 2 Hz.

291 Next, we examined synchronization for each trial. The sequence of tap times was converted into a
292 sequence of phases relative to the predicted pulse frequency, and the circular mean was computed for each
293 trial ((Batschelet, 1981), Figure 7b; blue circles). The grand mean was then computed for each complexity
294 level (Figure 7b, red line). The length, r, of the mean vector is the synchronization coefficient, a measure
295 of tapping variability (1 corresponds to no variability, and 0 corresponds to maximum variability), and the
296 angle, ϕ, of the mean vector indicates the relationship to the hypothetical pulse, where positive values of ϕ
297 indicate late tapping and negative values indicate early taps, on average (Batschelet, 1981).

298 At the lowest level of complexity, mean vectors clustered near the unit circle, indicating a high degree of
299 synchronization, and near a phase of 0, indicating in-phase tapping. As complexity increased, a growing
300 number of vectors clustered near zero, indicating lack of synchronization. The percentages of trials that
301 were synchronized at $r > 0.5$ (a highly conservative criterion Fisher, 1993) were 97%, 96%, 89%, 60%,
302 and 55% for complexity levels $0 - 4$, respectively. Thus, the more complex rhythms were more difficult
303 to synchronize with, however, the majority of the time participants synchronized at the predicted pulse
304 frequency, even for the rhythms that had no stimulus energy at that frequency. Finally, for the most
305 complex rhythms (levels 3 and 4) a significant number of mean relative phases clustered near a relative
306 phase of π, indicating anti-phase tapping. The proportion of anti-phase tapping changed across the five
307 levels of syncopation ($\chi^2(1, n = 565) = 64.59, p < .001$), but was equally distributed across tempos,
308 ($\chi^2(1, n = 565) = 3.16, p > .10$). Although the stimulus tempo range was relatively small, the fact that

309 this analysis showed that the majority of trials were well phase locked to the pulse indicates subjects were
310 not merely tapping near a preferred tempo (e.g., 2 Hz). If this had been the case, these trials would not
311 have shown such strong phase locking.

312 As predicted by the model, participants synchronized — either in-phase or anti-phase — predominantly
313 at the missing pulse frequency. This behavior is consistent with the prediction that formation of the
314 pulse percept arises due to entrainment of emergent neuronal oscillations. It also rules out the potential
315 alternatives that synchronization is merely a consequence of a common rhythmic input, or that the pulse
316 percept may arise due to linear resonance. Theoretically speaking, it is critical to distinguish the role of
317 a common stimulus frequency from the intrinsic dynamics of a emergent oscillation (Whittington et al.,
318 2000), and the missing pulse rhythms used here enabled us to dissociate the two. Such rhythms may also
319 be useful in studying the development of rhythm perception and synchronization (e.g., Kirschner and
320 Tomasello, 2009), as well as the capabilities and limitations of 'beat-deaf' individuals (e.g., Phillips-Silver
321 et al., 2011) and non-human animals (e.g., Patel et al., 2009; Cook et al., 2013; Large and Gray, in press),
322 as it allows us to tease apart the stimulus from any potential contribution of endogenous neural rhythms.
323 Additionally, more thorough analysis of synchronization behavior will enable us to fit model parameters
324 and evaluate the detailed predictions of the neurodynamic model.

325 Recent findings using the steady-state evoked potential (SS-EP) approach provide converging evidence for
326 the hypothesis of nonlinear resonance in pulse perception. In one study, meter imagery elicited subharmonic
327 resonance at metric frequencies for a stimulus with no energy at the to-be-imagined subharmonics
328 (Nozaradan et al., 2011). This study provided direct evidence of mode-locking, however, the subharmonic
329 resonance was not spontaneous, it was instructed. Therefore, it does not directly address the issue of
330 spontaneous pulse perception in complex rhythms. In another study, the amplitudes of the SS-EPs at pulse
331 and meter frequencies were shown to be selectively enhanced in syncopated rhythms even though acoustic
332 energy was not predominant at these frequencies (Nozaradan et al., 2012). While the result of that study
333 matched predictions for the sensory network (Large, 2010; Velasco and Large, 2011) such a finding is not
334 inconsistent with linear resonance (Oppenheim and Schafer, 1975), because no responses were found at
335 frequencies that were absent from the stimulus. This study demonstrated perceived pulse at frequencies
336 that were completely absent from the stimulus rhythms.

5 DISCUSSION

337 The hypothesis of neural resonance to musical rhythms has been supported by a great deal of behavioral
338 and electrophysiological evidence since Jones's original dynamic attending proposal nearly forty years
339 ago. Not only has the framework stood up to empirical testing, but it has turned out to be compatible
340 and complementary in interesting ways with sensory-motor theory, originally proposed by Todd as an
341 alternative conception of musical rhythm perception. It is not surprising that sensory and motor rhythms
342 interact, but what is somewhat unexpected is the possibility that oscillatory sensory-motor interaction itself
343 could give rise to the percepts of pulse and meter. These data do not speak to how the sensory and motor
344 systems interact; for that, additional work will be needed. Moreover, it remains to test the theory at the level
345 of neural population oscillations, using EEG and MEG in humans, which is, in fact, the level of prediction
346 at which the mathematical theory operates. Nevertheless, the fundamental prediction of the theory, that
347 pulse should be perceived in rhythms with no energy in the amplitude envelope at the pulse frequency, is
348 supported by behavioral evidence.

349   The phenomenon demonstrated empirically here, the perception of a "missing pulse," is reminiscent
350   in some ways of the "missing fundamental" phenomenon of pitch perception (for a review, see **?**). If the
351   energy at the fundamental frequency is removed from the complex spectrum of a periodic sound, the
352   perceived pitch remains unchanged, matching the pitch of a sinusoid with the frequency of the missing
353   fundamental (**??**). Both phenomena involve the perception of a frequency that is "missing" from the
354   stimulus. The mechanisms of pitch perception are still debated by theorists (**?**). However, the analogous
355   effect may indicate the involvement of a similar mechanism, i.e., neural oscillation, albeit at different time
356   scales. For example, a model based on the oscillations of cochlear nucleus chopper cells has been proposed
357   to account for complex pitch perception (**?**). It is important to note, however, that while these phenomena
358   are similar, they are not the same. Perceived pitch generally corresponds to the fundamental of a pitched
359   sound, whereas the perceived pulse of a musical rhythm is not its fundamental frequency.

360   A fundamental aspect of neural computation that our model does not explicitly take into account is
361   the discharge rate and the temporal codes of action potentials produced by single cells in the central
362   nervous system. It has been suggested, for example, that oscillatory signals themselves do not carry
363   information, but instead all information is represented in the spiking activity of cells (see, e.g., Shadlen
364   and Movshon, 1999). Recent neurophysiological experiments in monkeys have revealed that neurons in
365   SMA dynamically encode duration and serial order elements of the produced intervals in a synchronization-
366   continuation task (Merchant et al., 2013; Crowe et al., 2014; Merchant et al., 2015b). However, oscillations
367   and spikes interact in important ways, and efforts are currently being made to understand how spike
368   discharge information coordinates with local and global neural oscillatory activity (Buzsáki et al., 2012;
369   Kayser et al., 2009; Musacchia et al., 2013). For example, the active sensing framework of Schroeder and
370   colleagues has clear parallels with neural resonance theory that are beginning to be recognized and exploited
371   (Musacchia et al., 2013; Henry and Herrmann, 2013). This approach describes rhythmic attention in terms
372   of excitation-inhibition cycles tied to slow neural oscillations that can be entrained by rhythmic stimulation
373   (Lakatos et al., 2005, 2008; Schroeder et al., 2010). Entrainment of cortical rhythms is thought to engage
374   neurodynamic mechanisms of temporal prediction to segregate incoming information and organize spike
375   timing (Musacchia et al., 2013). Experiments are investigating concepts of attending dynamics at the
376   neuronal level and, in combination with the emerging study of behavioral synchronization in nonhuman
377   animals (Patel et al., 2009; Zarco et al., 2009; Honing et al., 2012; Cook et al., 2013; Large and Gray, in
378   press), promises insights that are capable of linking neuronal dynamics with perception, attention and
379   behavior.

380   Recently, Patel and Iversen have claimed that entrainment of neural oscillation to external rhythms is
381   not sufficient to explain beat perception (Patel and Iversen, 2014). In their view, neural oscillation cannot
382   account for the fact that: "*pure perception* of a musical beat (i.e., listening in the absence of overt movement)
383   strongly engages the motor system," and they suggest that "beat perception involves more than the passive
384   entrainment of neural responses to sound." They propose the *action simulation for auditory prediction*
385   (ASAP) hypothesis as an alternative. However, these objections represent misconceptions regarding the
386   fundamental nature of neural oscillation and the specific predictions of neural resonance theory. Here, we
387   have provided a specific model of pulse and meter perception based on oscillatory interactions between
388   auditory and motor systems, and we have tested the prediction that pulse perception is based on entrainment
389   of active, endogenous neural oscillations. Our result conclusively rules out passive synchronization of
390   neural activity with external rhythms. While the ASAP hypothesis makes predictions regarding the specific
391   neural structures involved in the perception of pulse and meter, it makes no predictions regarding the nature
392   of the computations that take place in these regions. By contrast, the neural resonance approach makes
393   specific predictions about the nature of the neurodynamic processes involved in musical rhythm perception

394 (see Large, 2008; Large and Snyder, 2009) and it explains a phenomenon that other models fail to explain.
395 It does not make strong commitments to specific brain regions, however, and more empirical data will be
396 necessary to understand how the various neural structures involved contribute to observed neurodynamics.
397 Nevertheless, the theory is based on a canonical model of neural oscillation, thus the predictions hold for
398 an entire family of dynamical systems, suggesting that it will scale as we learn more about the dynamical
399 interactions in the striato-thalamo-cortical network underlying rhythm perception. Neural resonance theory
400 is also consistent with Hebbian theory (Large, 2010), suggesting that it will be able to account for findings
401 regarding plasticity and perceptual narrowing in rhythmic enculturation (e.g., Hannon and Trehub, 2005).

402 Additionally, Patel and Iversen have identified neural resonance theory with Darwin's idea that "The
403 perception ... of musical ... rhythm is probably common to all animals, and no doubt depends on the
404 common physiological nature of their nervous systems" (Patel and Iversen, 2014). They argue against
405 neural resonance based on their interpretation that "[rhesus monkeys] do not perceive a beat in rhythmic
406 auditory patterns ... [and] could not learn to tap in synchrony with an auditory metronome". However,
407 their fundamental argument appears to be based on the assumption that if neural resonance underlies
408 pulse perception in humans, similar networks must be present in all animals. Moreover, Patel (2014) and
409 Fitch (2012) reason that if that is true, then all animals should be able to synchronize body movements to
410 rhythms. As we have argued elsewhere (Large and Gray, in press), species differences in perception-action
411 coordination do not speak to the question of whether or not the fundamental neural processes are oscillatory.
412 Species differences could be accommodated within a neural resonance framework, for example, simply in
413 terms of the differences in auditory-motor coupling (see, e.g., Merchant and Honing, 2013). Furthermore,
414 to the extent that ASAP is to be identified with Patel's "vocal learning hypothesis," recent observations of
415 entrainment to musical rhythms in a sea lion (Cook et al., 2013), and entrainment to simpler rhythms in a
416 chimpanzee (Hattori et al., 2013) and a bonobo (Large and Gray, in press) are providing mounting counter
417 evidence.

418 An understanding of the relationship between auditory neurophysiology, auditory population dynamics
419 and auditory perception is an elusive goal, due to the intricate circuitry, the many structural levels involved,
420 and the highly nonlinear nature of the neural responses. Significant theoretical advances will be necessary
421 to understand signal processing, pattern formation, and plasticity in this complex and highly nonlinear
422 system. Our approach draws upon the well-established theory of nonlinear dynamical systems to build
423 novel and increasingly comprehensive models, and make predictions for physiological and behavioral
424 experiments. This may allow us to directly link, for the first time, complex neuronal dynamics in animal
425 models, neural population dynamics observed in human EEG and MEG, and perceptual dynamics assessed
426 behaviorally, within an overarching theoretical framework that describes system dynamics at multiple
427 levels of observation.

REFERENCES

428 Ainsworth, M., Lee, S., Cunningham, M. O., Traub, R. D., Kopell, N. J., and Whittington, M. A. (2012).
429 Rates and rhythms: a synergistic view of frequency and temporal coding in neuronal networks. *Neuron*
430 75, 583

431 Aronson, D. G., Ermentrout, G. B., and Kopell, N. (1990). Amplitude response of coupled oscillators.
432 *Physica D: Nonlinear Phenomena* 41, 449

433 Barnes, R. and Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology* 41, 254–311

Bartolo, R. and Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *The Journal of Neuroscience* 35, 4635–4640

Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *The Journal of Neuroscience* 34, 3910–3923

Batschelet, E. (1981). *Circular Statistics in Biology* (London: Academic Press)

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., et al. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex* 45, 62–71

Berens, P. and Velasco, M. J. (2009). *CircStat for Matlab: Toolbox for circular statistics with Matlab*

Bergeson, T. R. and Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception* 23, 345–360

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology* 6, 145–238

Börgers, C. and Kopell, N. (2003). Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. *Neural Computation 15* 3, 509–539

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience,* 8, 183–208

Brunel, N. (2003). *Dynamics and plasticity of stimulus-selective persistent activity in cortical network models* (Cerebral Cortex)

Buzsáki, G. (2006). *Rhythms of the Brain* (New York: Oxford University Press)

Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currentseeg, ecog, lfp and spikes. *Nature Reviews - Neuroscience* 13, 407–420

Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304(5679), 1926–1929

Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S. J. A., Steinberg, F., and Large, E. W. (2010). Neural responses to complex auditory rhythms: the role of attending. *Front Psychol* 1, 224

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex* 18

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008b). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm lexity and musical training. *Journal of Cognitive Neuroscience* 20, 226–239

Cook, P., Rouse, A., Wilson, M., and Reichmuth, C. (2013). A California Sea Lion (Zalophus californianus) Can Keep the Beat: Motor Entrainment to Rhythmic Auditory Stimuli in a Non Vocal Mimic. *Journal of Comparative Psychology* 127, 412–27

Crowe, D. A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *The Journal of Neuroscience* 34, 11972–11983

Fisher, N. I. (1993). *Statistical Analysis of Circular Data* (Cambridge: Cambridge University Press)

Fitch, W. T. (2012). The biology and evolution of rhythm: Unravelling a paradox. In *Language and Music as Cognitive Systems*, eds. P. Rebuschat, M. Rohmeier, J. A. Hawkins, and I. Cross (Oxford University Press, Oxford, UK). 73–95

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *The Neurosciences and Music III: disorders and Plasticity. Annals of the New York Academy of Sciences, 2009* 1169, 89–92

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *The Journal of Neuroscience* 32, 1791–1802

479 Grahn, J. A. and Rowe, J. B. (2009). *Feeling the beat: premotor and striatal interactions in musicians and*
480   *non-musicians during beat processing* (Journal of Neuroscience)

481 Grahn, J. A. and Rowe, J. B. (2013). Finding and feeling the musical beat: striatal dissociations between
482   detection and prediction of regularity. *Cerebral cortex (New York, NY : 1991)* 23, 913–921

483 Hannon, E. E. and Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily
484   than adults. *Proceedings of the National Academy of Sciences of the United States of America* 102,
485   12289–12290

486 Hattori, Y., Tomonaga, M., and Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory
487   rhythm in a chimpanzee. *Sci. Rep.* 3

488 Henry, M. J. and Herrmann, B. (2013). *Low-Frequency Neural Oscillations Support Dynamic Attending in*
489   *Temporal Context*

490 Honing, H., Merchant, H., Háden, G., Prado, L., and Bartolo, R. (2012). Probing beat induction in Rhesus
491   monkeys: Is beat induction species-specific. In *Proceedings of the International Conference on Music*
492   *Perception and Cognition*. 454–455

493 Hoppensteadt, F. and Izhikevich, E. M. (1997). *Weakly connected neural networks*. Applied mathematical
494   sciences (New York: Springer)

495 Hoppensteadt, F. C. and Izhikevich, E. M. (1996a). Synaptic organizations and dynamical properties of
496   weakly connected neural oscillators I: Analysis canonical model. *Biological Cybernetics* 75, 117–127

497 Hoppensteadt, F. C. and Izhikevich, E. M. (1996b). Synaptic organizations and dynamical properties of
498   weakly connected neural oscillators II: Learning e information. *Biological Cybernetics* 75, 126–135

499 Hoppensteadt, F. C. and Izhikevich, E. M. (2005). Oscillatary neurocomputers with dynamic connectivity.
500   US Patent 6,957,204

501 Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates
502   early auditory responses. *Annals of the New York Academy of Sciences* 1169, 58–73

503 Izhikevich, E. M. (2000). *Neural excitability, spiking and bursting* (International Journal of Bifurcation
504   and Chaos)

505 Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory.
506   *Psychological Review* 83, 323–335

507 Jones, M. R. (2008). Musical time. In *Oxford Handbook of Music Psychology*, eds. S. Hallam, I. Cross,
508   and M. Thaut (Oxford University: Oxford). 81–92

509 Jones, M. R. and Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review* 96,
510   459–491

511 Jones, M. R., Johnston, H. M., and Puente, J. (2006). Effects of auditory pattern structure on anticipatory
512   and reactive attending. *Cognitive Psychology* 53, 59–96

513 Jones, M. R. and McAuley, J. D. (2005). Time judgments in global temporal contexts. *Perception &*
514   *Psychophysics* 67, 398–417

515 Jones, M. R. and Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of*
516   *Experimental Psychology: Human Perception & Performance* 23, 693–709

517 Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-phase coding boosts and
518   stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608

519 Kirschner, S. and Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in
520   preschool children. *Journal of Experimental d Psychology* 102, 299–314

521 Koepsell, K., Wang, X., Hirsch, J. A., and Sommer, F. T. (2010). Exploring the function of neural
522   oscillations in early sensory systems. *Front Neuroscience* 4, 53

523 Kung, S.-J., Chen, J. L., Zatorre, R. J., and Penhune, V. B. (2013). Interacting cortical and basal ganglia
524   networks underlying finding and tapping to the musical beat. *Journal of Cognitive Neuroscience* 25,
525   401–420

526 Laje, R. and Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent
527   neural networks. *Nature Neuroscience* 16, 925–936. doi:doi:10.1038/nn.3405

528 Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal
529   oscillations as a mechanism of attentional selection. *Science* 320, 110–113

530 Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory
531   hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of
532   Neurophysiology* 94, 1904–1911

533 Langner, G. (1992). Periodicity coding in the auditory system. *Hearing Research* 60, 115–142

534 Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In *The Psychology of Time*,
535   ed. S. Grondin (Emerald: Cambridge)

536 Large, E. W. (2010). Neurodynamics of music. In *Springer Handbook of Auditory Research: Music
537   Perception, M* (Springer: New York). 201–231

538 Large, E. W., Almonte, F., and Velasco, M. (2010). A canonical model for gradient frequency neural
539   networks. *Physica D: Nonlinear Phenomena* 239, 905–911

540 Large, E. W. and Gray, P. (in press). Spontaneous tempo and entrainment in a bonobo (pan paniscus).
541   *Journal of Comparative Psychology*

542 Large, E. W. and Jones, M. R. (1999). The dynamics of attending: How people track time varying events.
543   *Psychological Review* 106, 119–159

544 Large, E. W. and Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York
545   Academy of Sciences* 1169, 46–57. doi:10.1111/j.1749-6632.2009.04550.x

546 Ledoux, E. and Brunel, N. (2011). Dynamics of networks of excitatory and inhibitory neurons in response
547   to time-dependent inputs. *Frontiers in Computational Neuroscience* 5.

548 Lerdahl, F. and Jackendoff, R. S. (1983). *A Generative Theory of Tonal Music* (Cambridge: MIT Press)

549 London, J. M. (2004). *Hearing in Time: Psychological aspects of musical meter* (New York, New York:
550   Oxford University Press)

551 McAuley, J. D. and Kidd, G. R. (1995). Temporally directed attending in the discrimination of tempo:
552   Further evidence for an entrainment model. *Journal of Acoustical Society of America* 97, 3278

553 Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., and Fitch, W. T. (2015a). Finding the beat: a neural
554   perspective across humans and non-human primates. *Phil. Trans. R. Soc. B* 370

555 Merchant, H. and Honing, H. (2013). Are non-human primates capable of rhythmic entrainment? Evidence
556   for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience* 7, 274

557 Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., et al. (2015b). Sensorimotor
558   neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *European
559   Journal of Neuroscience* 41, 586–602. doi:doi:10.1111/ejn.12811

560 Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval tuning in the primate medial premotor
561   cortex as a general timing mechanism. *The Journal of Neuroscience* 33, 9082–9096

562 Morillon, B., Hackett, T. A., Kajikawa, Y., and Schroeder, C. E. (2015). Predictive motor control of sensory
563   dynamics in auditory active sensing. *Current Opinion in Neurobiology* 31, 230–238

564 Morillon, B., Schroeder, C. E., and Wyart, V. (2014). Motor contributions to the temporal precision of
565   auditory attention. *Nature communications* doi:DOI:10.1038/ncomms6255

566 Murdock, J. (2003). *Normal Forms and Unfoldings for Local Dynamical Systems*. Springer Monographs
567   in Mathematics (Springer)

568 Musacchia, G., Large, E. W., and Schroeder, C. E. (2013). *Thalamocortical mechanisms for integrating*
569 *musical tone and rhythm* (Hearing Research)

570 Nozaradan, S., Chemin, B., and Mouraux, A. (2013). *From sounds to movement and back: how movement*
571 *shapes internal representation of musical rhythms* (Frontiers in human neuroscience)

572 Nozaradan, S., Peretz, I., Missal, M., and Mouraux, A. (2011). Tagging the neuronal entrainment to beat
573 and meter. *Journal of Neuroscience* 31, 10234–40

574 Nozaradan, S., Peretz, I., and Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter
575 embedded in a musical rhythm. *Journal of Neuroscience* 32, 17572–17581

576 Oppenheim, A. V. and Schafer, R. W. (1975). *Digital Signal Processing* (Englewood Cliffs, NJ: Prentice
577 Hall)

578 Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was darwin wrong? *PLoS Biol* 12,
579 e1001821. doi:10.1371/journal.pbio.1001821

580 Patel, A. D. and Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the
581 action simulation for auditory prediction (asap) hypothesis. *Frontiers in Systems Neuroscience* 8.
582 doi:10.3389/fnsys.2014.00057

583 Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Experimental evidence for
584 synchronization to a musical beat in a nonhuman animal. *Current Biology* 19, 830

585 Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., et al. (2011). Born to
586 dance but beat deaf: a new form of congenital amusia. *Neuropsychologia* 49, 961–969

587 Phillips-Silver, J. and Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception.
588 *Science* 308, 1430–1430

589 Phillips-Silver, J. and Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic
590 movement. *Cognition* 105, 546

591 Rauschecker, J. P. and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates
592 illuminate human speech processing. *Nature Neuroscience* 12, 718–724

593 Repp, B. H. (2008). *Multiple temporal references in sensorimotor synchronization with metrical auditory*
594 *sequences* (Psychological Research)

595 Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of Active
596 Sensing and perceptual selection. *Current Opinion in Neurobiology* 20, 172–176

597 Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework.
598 *TRENDS in Cognitive Sciences* 11

599 Shadlen, M. N. and Movshon, J. A. (1999). Synchrony unbound: A critical evaluation of the temporal
600 binding hypothesis. *Neuron* 24, 67–77

601 Slézia, A., Hangya, B., Ulbert, I., and Acsády, L. (2011). Phase Advancement and Nucleus-Specific Timing
602 of Thalamocortical Activity during Slow Cortical Oscillation. *The Journal of Neuroscience* 31, 607–617

603 Snyder, J. S. and Large, E. W. (2005). *Gamma-band activity reflects the metric structure of rhythmic tone*
604 *sequences* (Cognitive Brain Research)

605 Stefanescu, R. A. and Jirsa, V. K. (2008). A low dimensional description of globally coupled heterogeneous
606 neural networks of excitatory and inhibitory neurons. *PLoS Computational Biology* 4, e1000219

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

610 Steriade, M., McCormick, D. A., and Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping
611 and aroused brain. *Science* 262, 685

612 Sussillo, D. and Abbott, L. (2009). Generating coherent patterns of activity from chaotic neural networks.
613   *Neuron* 63, 544–557

614 Todd, N. P. M. (1994). The auditory primal sketch: A multi-scale model of rhythmic grouping. *Journal of*
615   *New Music Research* 23, 25–69

616 Todd, N. P. M. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*
617   97, 1940–1949

618 Varela, F., Lachaux, J.-P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: Phase synchronization
619   and large-scale integration. *Nature Review Neuroscience* 2, 229–239

620 Velasco, M. J. and Large, E. W. (2011). Pulse detection in syncopating rhythms using neural oscillators. In
621   *Proceedings of the 12th Annual Conference he International Society for Music Information Retrieval*

622 Vos, P. G. M. M. (1973). *Waarneming van metrische toonreeksen* (Nikmegen: Stichting Studentenpers)

623 Whittington, M. A., Traub, R. D., Kopell, N., Ermentrout, G. B., and Buhl, E. H. (2000). Inhibition-based
624   rhythms: experimental and mathematical observations on network dynamics. *International Journal*
625   *Psychophysiology* 38, 315–336

626 Will, U. and Berg, E. (2007). Brain wave synchronization and entrainment to periodic acoustic stimuli.
627   *Neuroscience Letters* 424, 55–60

628 Wilson, H. R. and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and
629   thalamic nervous tissue. *Kybernetik* 13, 5–80

630 Wing, A. M. and Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses.
631   *Perception and Psychophysics* 14, 5–12

632 Wing, A. M. and Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Perception and*
633   *Psychophysics* 13, 455–460

634 Yee, W., Holleran, S., and Jones, M. R. (1994). Sensitivity to event timing in regular and irregular
635   sequences: Influences of musical skill. *Perception & Psychophysics* 56, 461–471

636 Zarco, W., Merchant, H., Prado, L., and Mendez, J. C. (2009). Subsecond timing in primates: comparison
637   of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology* 102,
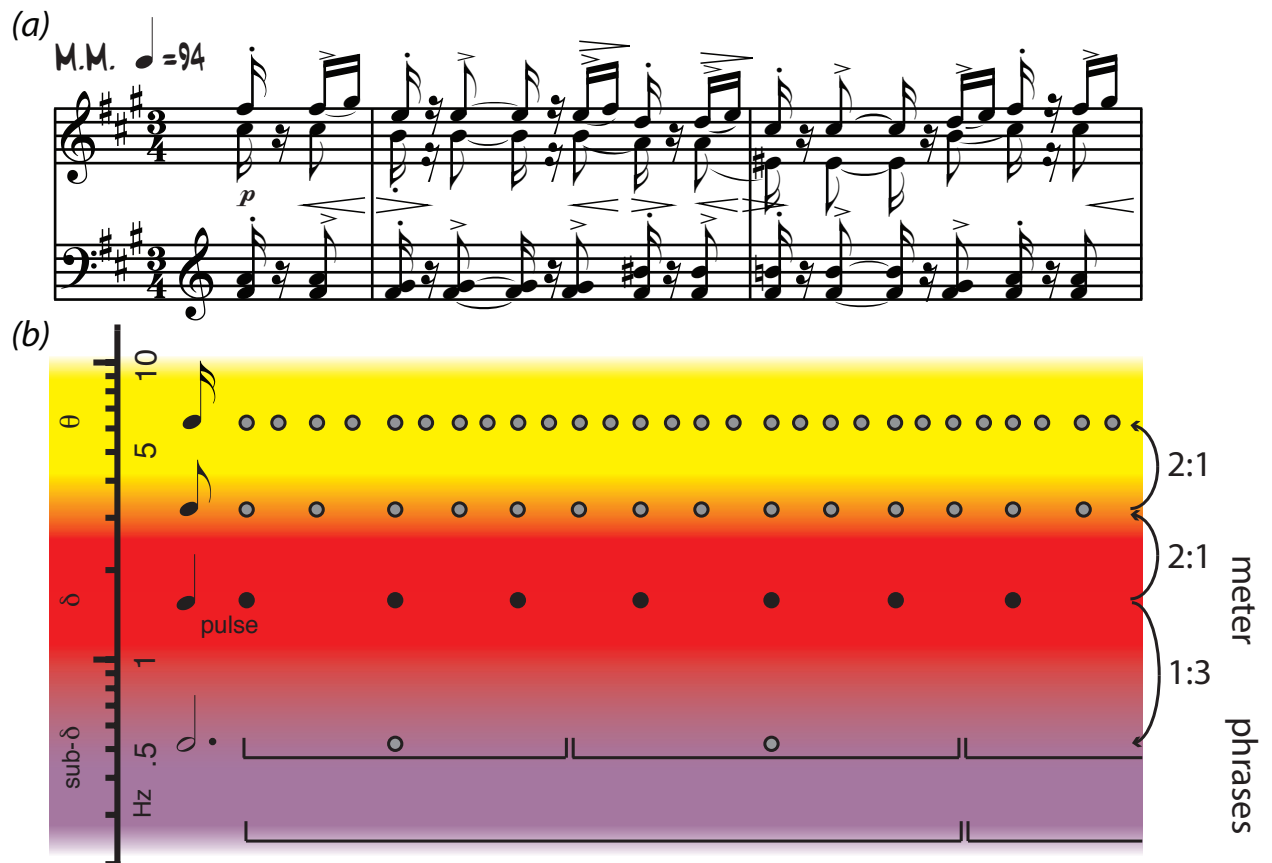638   3191–202

FIGURES

Figure 1: (a) Piano score of Isaac Albeniz's Iberia II, Triana. (b) Annotation showing pulse (black dots), metrical structure (all dots), and phrasing (brackets). Pulse frequencies overlap with cortical oscillations in the delta range, while other metrical frequencies extend into theta and sub-delta ranges. Frequency relationships among metrical levels include harmonics, subharmonics, and other integer ratios. Adapted from Large, 2014.
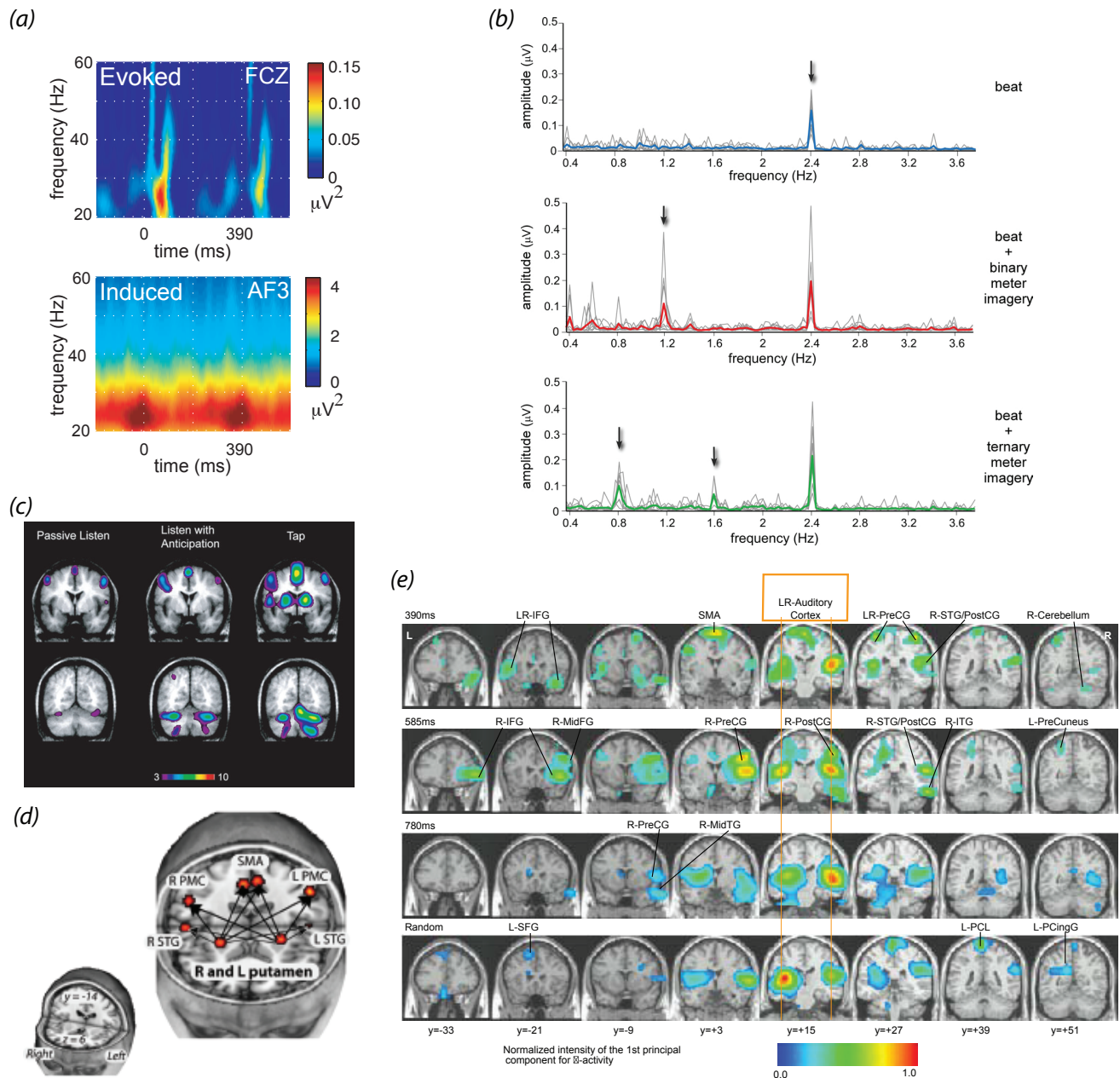
Figure 2: (a) EEG revealed synchronized fluctuations in induced beta- and gamma-band power that anticipated tone onsets, were sensitive to intensity accents, and persisted when expected tones were omitted. Evoked activity occurred after tone onsets and were strongly diminished after tone omissions (Snyder and Large, 2005) . (b) A periodic rhythm elicited a steady-state evoked potential (SS-EP) at the stimulus repetition frequency, and meter imagery elicited subharmonic resonances corresponding to the metric interpretation of this periodic rhythm (Nozaradan et al., 2011). (c) fMRI showed that listening to musical rhythms recruits both auditory and motor areas of the brain even in perception tasks without a motor component (Chen et al., 2008a). (d) Functional connectivity analysis revealed a cortico-subcortical network including the putamen, SMA, and PMC under conditions that may require internal generation of the pulse (Grahn and Rowe, 2009). (e) MEG revealed oscillatory interactions in a striato-thalamo-cortical network (Fujioka et al., 2012).
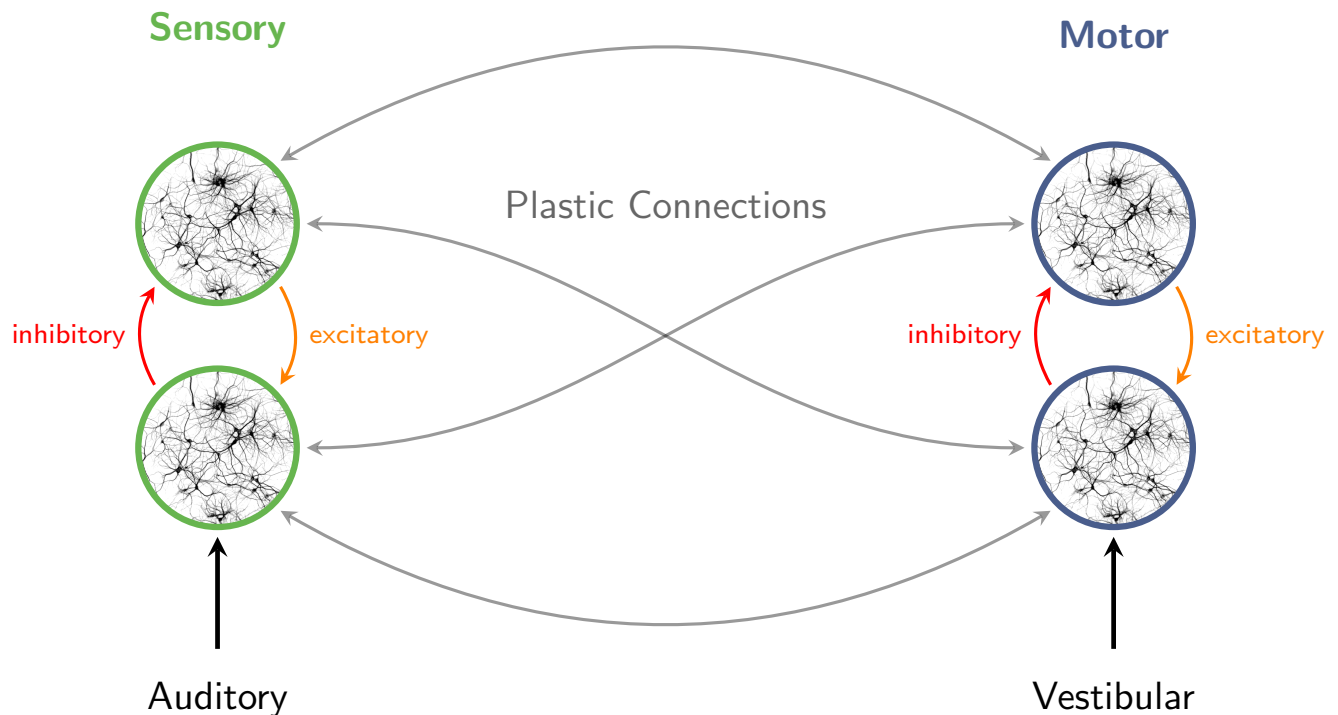
Figure 3: Interaction of excitatory and inhibitory neuronal populations gives rise to population rhythms in sensory and motor networks. When sensory stimuli are presented in a periodic pattern, auditory cortical oscillations entrain to the structure of the stimulus stream. The perception of pulse and meter in music also involves broadly distributed motor systems. Auditory-motor coupling is reciprocal, as vestibular stimulation can influence auditory rhythm perception. Connections within and between sensory and motor systems are assumed to be plastic.
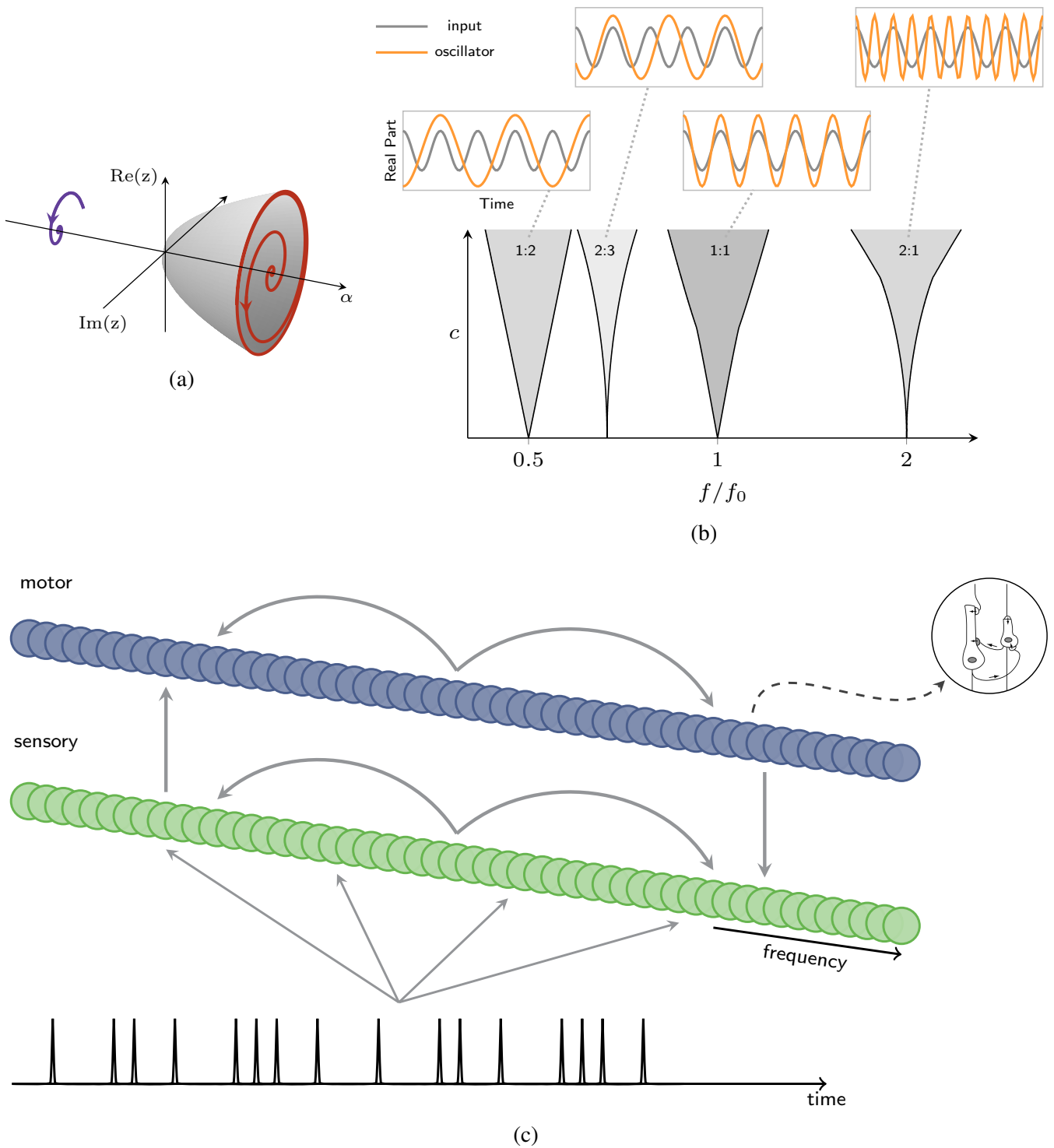
Figure 4: (a) Intrinsic oscillatory dynamics can take the form of a Hopf bifurcation, where a value $\alpha = 0$ is the critical point, between damped (left) and spontaneous (right) oscillation. (b) Mode locking in the canonical model. Within each resonance region (shaded), the canonical model mode-locks to input at the ratio shown in figure (c: coupling strength, f: oscillator's intrinsic frequency, f_0: input frequency). Insets show the inputs and traces produced by a canonical model. (c) Architecture of a model that captures interacting oscillatory dynamics in sensory and motor networks. A rhythm is input to a sensory network; sensory and motor networks are reciprocally connected, providing input to one another.
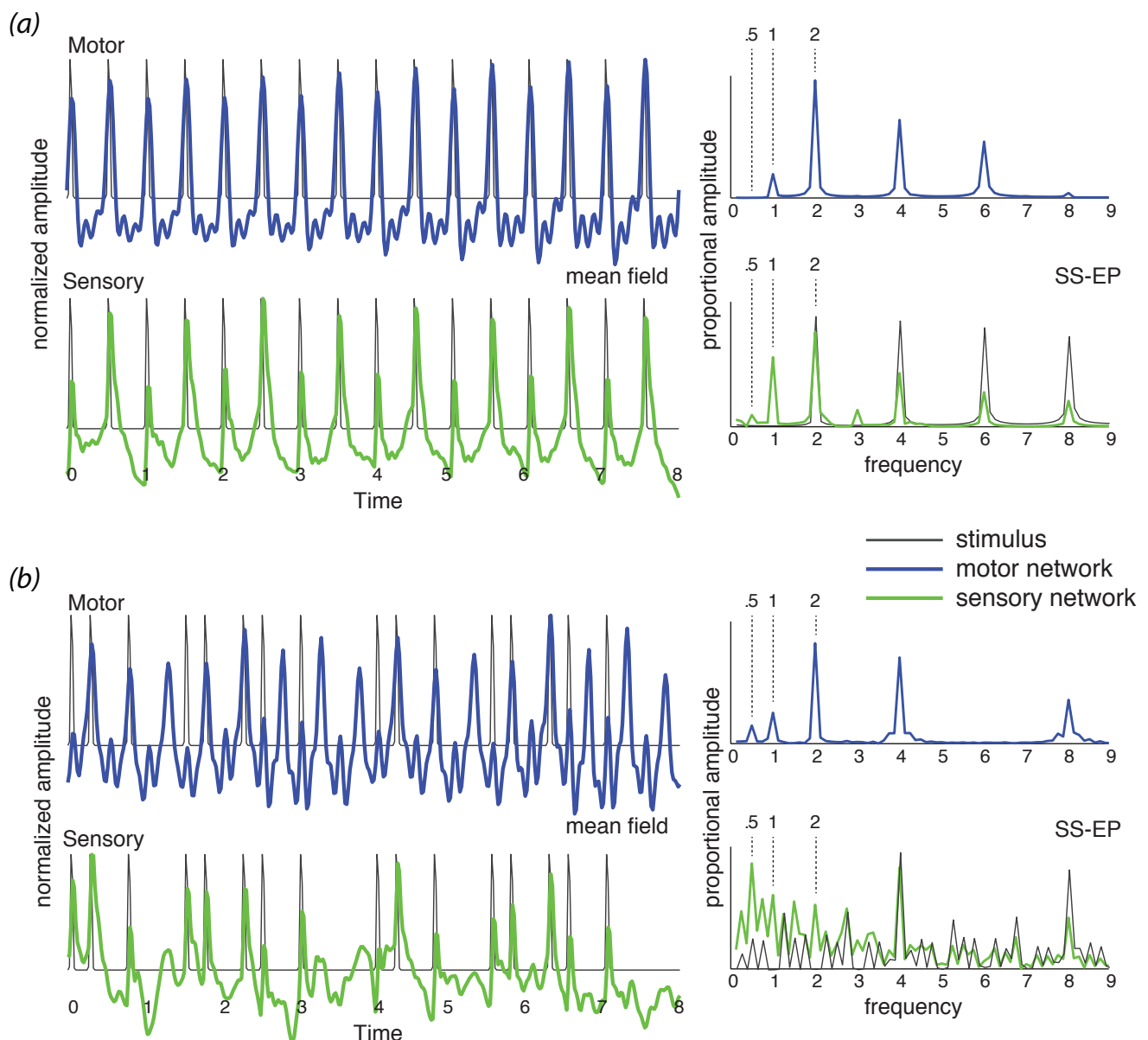
Figure 5: The network of Figure 4c was stimulated with (a) an isochronous rhythm and (b) a "missing pulse" rhythm. Output of the sensory network is in green and output of the motor network is in blue. The mean field time series (left) was obtained by summing the output of all the oscillators in the network over time. The SS-EP is obtained by Fourier analysis (DFT) of the mean field time series. The Fourier analysis (DFT) of the stimulus envelope is shown in black on the sensory SS-EP axis.
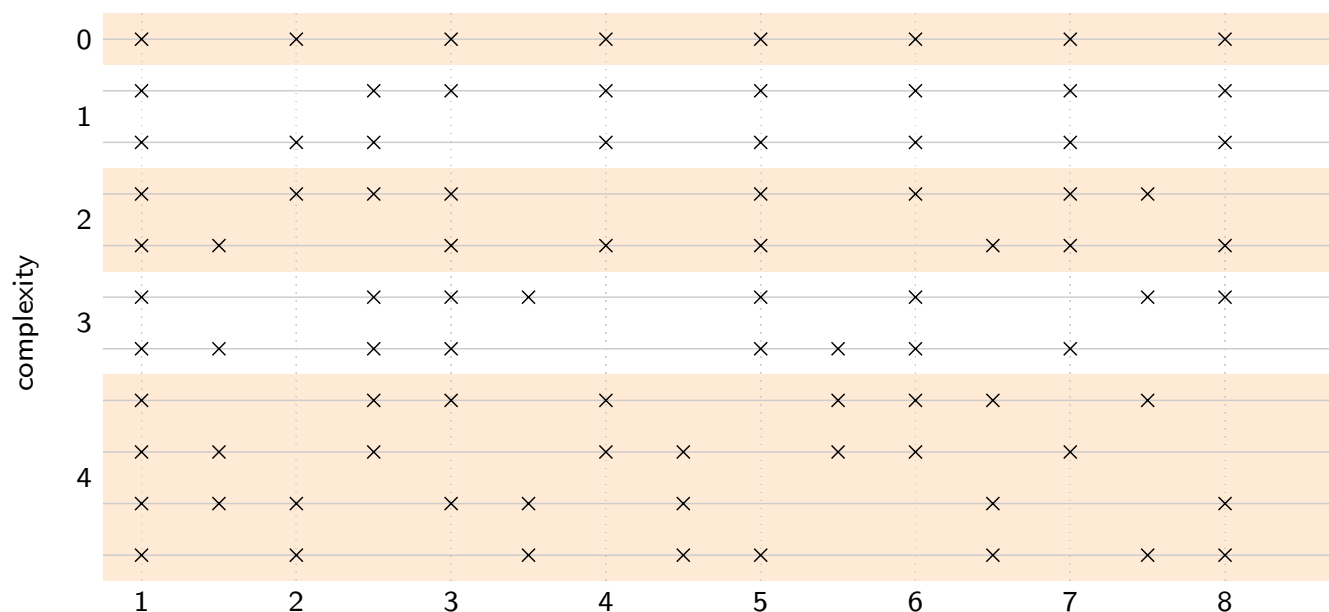
Figure 6: Each stimulus pattern contained eight events distributed over eight pulse cycles. Complexity was varied by manipulating the number of events that fell in phase versus anti-phase with the intended pulse. At complexity level 0, all eight events were in-phase, at complexity level 1, one event was anti-phase, at complexity level 2, two events were anti-phase, and so on. Level 4 was the most complex, with four events in-phase and four events anti-phase. One isochronous control (level 0), two level 1 patterns, two level 2 patterns, two level 3 patterns, and four level 4 patterns were used, for a total of 11 rhythms.
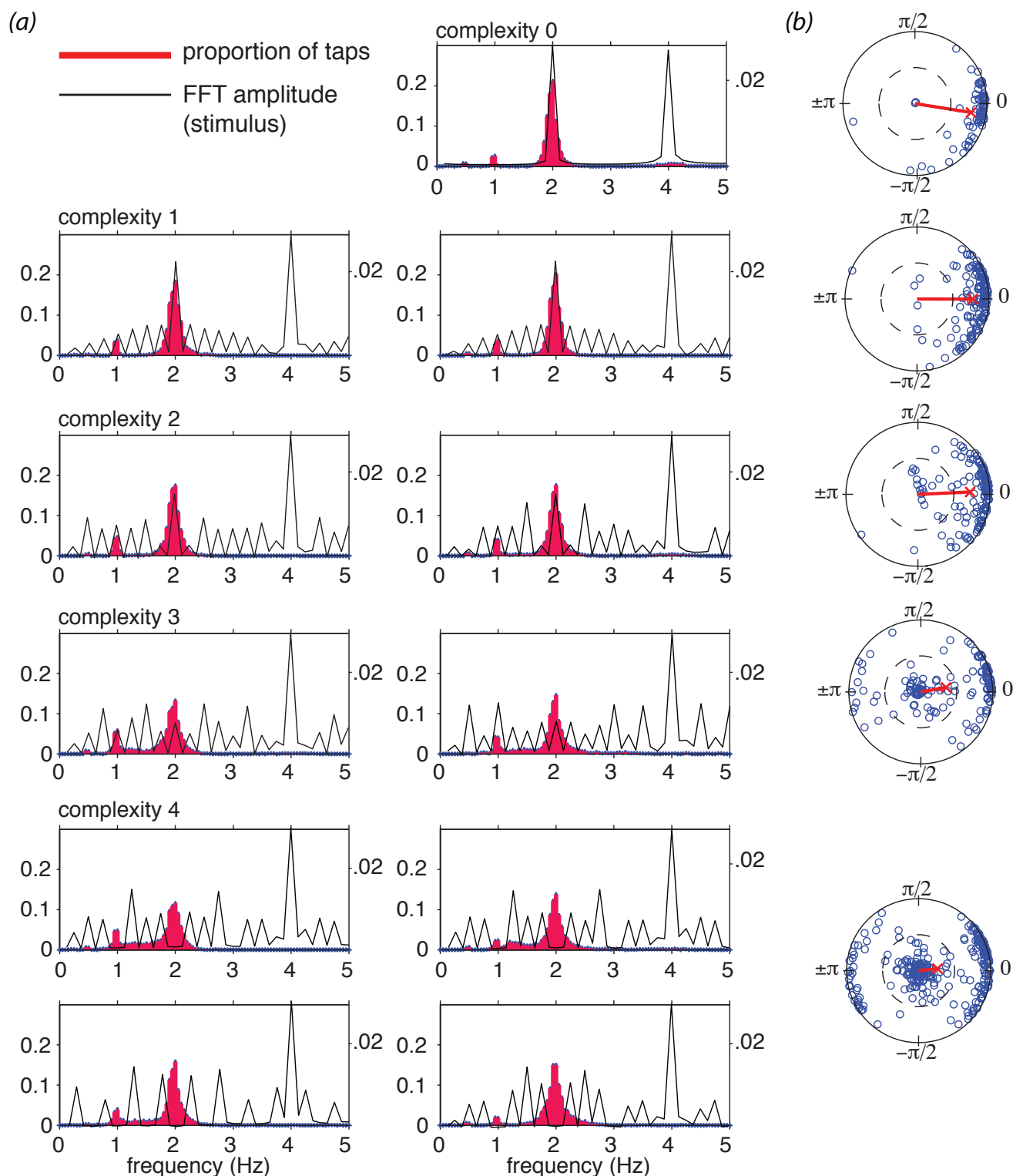
Figure 7: (a) Tapping frequencies were normalized to a 2 Hz (120bpm) tempo to allow comparison between trials at different tempos. Tapping frequency distributions (red histograms) were computed by binning normalized instantaneous tapping frequencies from 0 Hz to 5.00 Hz in bins widths of 0.05 Hz. Distributions were computed for each rhythm separately, including every tap interval across trials. Black lines show amplitude spectrum of the stimulus envelope for comparison. (b) Circular means of tap phases for each trial (blue circles) and grand mean for each complexity level (red line).