Differential contributions of synaptic and intrinsic inhibitory currents to speech segmentation via flexible phase-locking in neural oscillators

Benjamin R. Pittman-Polletta[1*], **Yangyang Wang**[2], **David A. Stanley**[1], **Charles E. Schroeder**[3], **Miles A. Whittington**[4], **Nancy J. Kopell**[1]

***For correspondence:**
benpolletta@gmail.com (BRPP)

[1]Department of Mathematics & Statistics, Boston University, Boston, MA; [2]Department of Mathematics, University of Iowa, Iowa City, IA; [3]Translational Neuroscience Division, Center for Biomedical Imaging and Neuromodulation, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA; [4]Hull York Medical School, York, UK

Abstract Hypotheses suggest speech segmentation is executed by a hierarchy of oscillators in auditory cortex, with theta (3-7 Hz) rhythms playing a key role by phase-locking to syllable boundaries. Reliable synchronization to quasi-rhythmic inputs, whose variable frequency can dip below cortical theta frequencies (down to 1 Hz), requires “flexible" theta oscillators, whose neural implementation remains unknown. Using biophysical computational models, we found that the flexibility of phase-locking depends on the types of hyperpolarizing currents that pace neural oscillators. Simulated cortical theta oscillators flexibly phase-locked to slow inputs when these inputs caused both (i) spiking and (ii) the subsequent buildup of outward current sufficient to delay further spiking until the next input. The greatest flexibility in phase-locking arose from a synergistic interaction between intrinsic currents, not replicated by synaptic currents at similar timescales. Our results suggest synaptic and intrinsic inhibition contribute to regular and flexible phase-locking in neural oscillators, respectively.

Introduction

How the cortex derives robust representations of linguistic objects remains a challenging problem. Relying on both temporally proximal and distal cues about speech rate and phonology, the brain decodes the speech stream in parallel at multiple levels of abstraction, producing a hierarchy of phonemes, syllables, words, and phrases (***Christiansen and Chater, 2016***), and predicting the timing and salience of upcoming speech (***Dilley and Pitt, 2010***; ***Dilley et al., 2010***; ***Brown et al., 2011***; ***Baese-Berk et al., 2014***; ***Brown et al., 2015***). Psychophysical and neurophysiological evidence suggests that this sampling, parsing, and prediction rely in part on a hierarchy of brain rhythms (***Ahissar et al., 2001***; ***Luo and Poeppel, 2007***; ***Nourski et al., 2009***; ***Hertrich et al., 2012***; ***Peelle et al., 2012***; ***Doelling et al., 2014***; ***Ding et al., 2016***; ***Riecke et al., 2017***; ***Zoefel et al., 2018***) – periodic fluctuations in the activity of neuronal populations (***Buzsaki, 2006***) – that mirror and align both with the temporal structure of speech and with each other (***Lakatos et al., 2005***; ***Schroeder and Lakatos, 2009***; ***Gross et al., 2013***; ***Henry et al., 2014***; ***Mai et al., 2016***; ***Pefkou et al., 2017***).

Conventional models of speech processing (***Marslen-Wilson, 1987***; ***Luce and CONOR, 2005***; ***Stevens, 2005***) suggest that decoding proceeds by matching chunks of speech of different durations with stored linguistic memory patterns or templates. Recent oscillation-based models have postulated that this template-matching is facilitated by a preliminary segmentation step (***Ghitza, 2011***; ***Giraud and Poeppel, 2012***; ***Ghitza, 2016***), which determines candidate speech segments for template matching, in the process tracking speech speed and allowing the adjustment (within limits) of sampling and segmentation rates (***Bosker and Ghitza, 2018***; ***Penn et al., 2018***); segmentation plays a key role in explaining a range of counterintuitive psychophysical data that challenge conventional models of speech perception (***Ghitza and Greenberg, 2009***; ***Ghitza, 2012***, ***2014***, ***2016***). A number of conceptual hypotheses (***Schroeder et al., 2008***; ***Ghitza, 2011***; ***Giraud and Poeppel, 2012***; ***Arnal and Giraud, 2012***; ***Ghitza, 2013***; ***Lewis and Bastiaansen, 2015***; ***Morillon and Schroeder, 2015***) suggest that ongoing cortical rhythms, entrained to the regular acoustic features of the speech stream that indicate the boundaries between linguistic units (***Rosen, 1992***; ***Hirst and Di Cristo, 1998***; ***Yang, 2007***; ***Yang et al., 2014***), effect this preliminary grouping of auditory input into a hierarchy of segments, each of which is subject to further phonemic, syntactic, and semantic processing. An important point (***Schroeder et al., 2008***) is that the frequencies of endogenous rhythms observed in auditory cortex – which include δ (∼1-4 Hz), θ (∼4-8 Hz), and β/γ (∼15-60 Hz) frequencies (***Lakatos et al., 2005***) – have timescales that mirror the timescales of linguistic units – namely, words and phrases (∼250-2000 ms), syllables (∼100-250 ms), and phonemes (∼20-100 ms, ***Selkirk*** (***1980***)). The hierarchical organization of auditory cortical brain rhythms, with β/γ rhythms nested within certain phases of the θ cycle, and θ rhythms in turn nested within certain phases of the δ cycle (***Lakatos et al., 2005***), makes rhythm phase information consistent with the hierarchy of linguistic segmentation.

While speech is a multiscale phenomenon, the modulation spectrum of continuous speech is dominated by syllabic rate amplitude fluctuations (***Ohala, 1975***; ***Chandrasekaran et al., 2009***; ***Elliott and Theunissen, 2009***; ***Ding et al., 2017***), restricted by the motor physiology of the speech apparatus to δ/θ frequencies (∼1-9 Hz, ***Ohala*** (***1975***); ***Chandrasekaran et al.*** (***2009***); ***Elliott and Theunissen*** (***2009***); ***Ding et al.*** (***2017***)), and critical for speech comprehension (***Elliott and Theunissen, 2009***; ***Ghitza, 2012***; ***Drullman et al., 1994***; ***Miller and Licklider, 1950***; ***Huggins, 1964***; ***Stilp et al., 2010***; ***Ghitza and Greenberg, 2009***). Syllabic frequencies, especially in the θ range, are central in auditory processing, as well: attentional entrainment to auditory rhythms operates best over the δ/θ range (***Lakatos et al., 2013***), perhaps supported by an active segmentation mechanism that operates on a timescale of ∼140-250 ms (∼4-7 Hz) to process sound input in syllable-sized chunks (***Teng et al., 2017***), and θ-frequency speech-brain entrainment is particularly relevant for speech processing (***Ahissar et al., 2001***; ***Luo and Poeppel, 2007***; ***Nourski et al., 2009***; ***Hertrich et al., 2012***; ***Peelle et al., 2012***; ***Doelling et al., 2014***; ***Ding et al., 2016***), with recent experiments suggesting a causal role (***Riecke et al., 2017***; ***Wilsch et al., 2017***, ***2018***; ***Zoefel et al., 2018***). Cortical θ rhythms – especially prominent in the spontaneous activity of primate auditory cortex (***Lakatos et al., 2005***) – seem to perform an essential function in syllable segmentation (***Ghitza and Greenberg, 2009***; ***Ghitza, 2014***, ***2012***; ***Doelling et al., 2014***) by marking the high energy θ-timescale features of speech as putative syllable boundaries (***Ghitza, 2011***, ***2012***; ***Hyafil et al., 2015***; ***Ten Oever and Sack, 2015***). Incorporating the centrality of the θ rhythm as “master" of the oscillatory hierarchy, seminal functional (***Ghitza and Greenberg, 2009***) and computational (***Hyafil et al., 2015***; ***Räsänen et al., 2018***) models have proposed that putative syllables segmented by θ rhythmic circuits are encoded by spiking in γ oscillatory circuits, while δ-rhythmic circuits may overlay δ-timescale prosodic information on this syllabic parse (***Ghitza, 2017***).

Syllable lengths vary over syllables, speakers, and languages, within a restricted range of “acceptable" syllable lengths (***Ghitza, 2014***). This variability places particular demands on the cortical θ oscillators tasked with syllabic segmentation. To track syllable boundaries occurring quasi-rhythmically, i.e. at variable intervals, the auditory cortical θ oscillator must be “flexible" – able to lock, cycle-by-cycle, to a quasi-rhythmic input with a broad range of instantaneous frequencies, including frequencies below the oscillator’s intrinsic frequency (***Ghitza, 2011***, ***2012***). The canonical implementation of this flexible oscillator is a voltage controlled oscillator in a phase-locked loop;

the mechanisms by which neural circuits implement oscillators exhibiting flexibility in phase-locking remain largely unexplored. If existing phenomenological models provided a complete explanation of this capability, then any θ oscillator should be able to perform this function. We show here, however, that the subtleties of the biophysical mechanisms giving rise to cortical θ oscillations can make a difference in their flexibility.

The major functional implication that concerns us is segmentation. For our purposes, segmentation refers to the ability of an oscillator to produce an output *only* during certain (high-energy) segments of a periodic or quasi-rhythmic input; and furthermore to produce output during *every* such segment. These high-energy segments represent syllable boundaries (occurring during the high-energy vocalic portion of the syllable). Toward this end, we explored the ability of biophysical computational models of neural θ oscillators to exhibit phase-locked spiking to strong periodic and quasi-rhythmic inputs having a range of input frequencies. Our θ oscillators were paced by (i) θ-timescale synaptic inhibition, or (ii) θ-timescale subthreshold oscillations (STOs) resulting from intrinsic θ-timescale hyperpolarizing currents, or (iii) both. Half also included intrinsic "super-slow" (δ-timescale) hyperpolarizing currents. While much is known about phase-locking in neural oscillators (***Ermentrout, 1981***, ***1996***; ***Kopell and Ermentrout, 2002***; ***Achuthan and Canavier, 2009***; ***Canavier and Achuthan, 2010***), few studies have examined the strong forcing regime, in which input pulses are strong enough to elicit spiking; little is known about how oscillator parameters influence phase-locking to inputs much slower or faster than an oscillator's intrinsic frequency; and few published studies explore oscillators exhibiting intrinsic outward currents on multiple timescales.

While the ability of our models to phase-lock to inputs faster than their intrinsic frequency was uniformly high and dependent on input strength, our results suggest that different types of inhibitory currents dramatically change the ability of θ oscillators to phase-lock to inputs slower than their intrinsic frequency over a range of input strengths. The central principle revealed by our models is that frequency flexibility is enabled by a buildup of outward (inhibitory) current during each input, which is sufficiently long-lasting to silence spiking during the period between successive inputs. Super-slow currents, having a timescale slower than that of intrinsic periodic spiking, enabled this buildup for oscillators paced by either synaptic inhibition or STOs. However, STOs interacted with these adapting currents synergistically to enable extremely flexible spike phase-locking that was absent from oscillators paced by synaptic inhibition. Thus, while inhibition-based oscillators phase-locked to rhythms within a relatively restricted frequency range (***Cannon and Kopell, 2015***; ***Sherfey et al., 2018a***), θ oscillators whose dynamics arose from a complex interplay of intrinsic currents in single cells (***Carracedo et al., 2013***) exhibited more flexible phase-locking. The features that facilitate flexible phase-locking come together in a single-compartment Hodgkin-Huxley model, reproducing *in vitro* data from layer 5 pyramidal cells with a θ-timescale resonance and a unique pattern of δ-nested θ-rhythmic spiking (***Carracedo et al., 2013***). Our computational model of these cells exhibited spiking entrainment to periodic input pulses with frequencies ranging over a broad interval, as well as to more realistic quasi-rhythmic inputs. The pattern of spike phase-locking observed in this oscillator – resulting from a buildup of outward current over the time course of each input – may contribute to the ability of auditory circuits to effectively segment and parse quasi-rhythmic signals, such as speech.

Methods

All simulations were run on the MATLAB-based programming platform DynaSim (***Sherfey et al., 2018b***), a framework specifically designed by our lab for efficiently prototyping, running, and analyzing simulations of large systems of coupled ordinary differential equations, enabling in particular evaluation of their dynamics over large regions of parameter space. DynaSim is open-source and all models will be made publicly available using this platform.

Table 1. Currents.

I_{Na} & $I_{\text{Na,SOM}}$	$g_{\text{Na}} m_{\text{Na}}^3 h \left(V - E_{\text{Na}}\right)$
I_{KDR} & $I_{\text{KDR,SOM}}$	$g_{\text{KDR}} m_{\text{KDR}}^4 \left(V - E_{\text{K}}\right)$
I_{leak} & $I_{\text{leak,SOM}}$	$g_{\text{leak}} \left(V - E_{\text{leak}}\right)$
I_{m}	$g_{\text{m}} n \left(V - E_{\text{K}}\right)$
I_{Nap}	$g_{\text{Nap}} m_{\text{Nap}} \left(V - E_{\text{Nap}}\right)$
I_{Ca}	$g_{\text{Ca}} s^2 \left(V - E_{\text{Ca}}\right)$
I_{KSS}	$g_{\text{KSS}} q \left(V - E_{\text{K}}\right)$
I_{inh} & I_{exc}	$g_{\text{pre}\to\text{post}} s_{\text{pre}\to\text{post}} \left(V_{\text{post}} - E_{\text{pre}\to\text{post}}\right)$

Table 2. Equilibrium Voltages.

	RS	FS
E_{Na}	40	50
E_{K}	-80	-95
E_{leak}	-65	-70
E_{Nap}	50	–
E_{Ca}	120	–
$E_{\text{RS}\to\text{FS}}$		0
$E_{\text{FS}\to\text{RS}}$	-95	

Model equations

Our models consisted of at most two cells, a regular spiking (RS) pyramidal cell and an inhibitory interneuron with a timescale of inhibition like that observed in somatostatin-positive interneurons (SOM). Each cell was modeled as a single compartment with Hodgkin-Huxley dynamics. In our RS model, the membrane currents consisted of fast sodium (I_{Na}), delayed-rectifier potassium (I_{KDR}), leak (I_{leak}), slow potassium or m- (I_{m}), and persistent sodium (I_{Nap}) currents taken from a model of a guinea-pig cortical neuron (***Gutfreund et al., 1995***), and calcium (I_{Ca}) and super-slow potassium (I_{KSS}, calcium-activated potassium in this case) currents with dynamics from a hippocampal model (***Traub et al., 1991***). The voltage $V(t)$ was given by the equation

$$C\frac{dV}{dt} = I_{\text{app}} - I_{\text{Na}} - I_{\text{KDR}} - I_{\text{leak}} - I_{\text{m}} - I_{\text{Nap}} - I_{\text{Ca}} - I_{\text{KSS}} - I_{\text{inh}}$$

where the capacitance $C = 2.7$ reflected the large size of deep-layer cortical pyramidal cells, and I_{app}, the applied current, was given by

$$I_{\text{app}}(t) = g_{\text{app}} \left[\left(\frac{t}{\tau_{\text{trans}}} \chi_{\{t \leq \tau_{\text{trans}}\}}(t) + \chi_{\{t > \tau_{\text{trans}}\}}(t) \right) + p_{\text{noise}} W(t) \right]$$

with the transition time $\tau_{\text{trans}} = 500$ ms, the noise proportion $p_{\text{noise}} = 0.25$, and $W(t)$ a white noise process. For SOM cells, the membrane currents consisted of fast sodium ($I_{\text{Na,SOM}}$), delayed-rectifier potassium ($I_{\text{KDR,SOM}}$), and leak ($I_{\text{leak,SOM}}$) currents. The voltage $V(t)$ was given by the equation

$$C_{\text{SOM}}\frac{dV}{dt} = I_{\text{app,SOM}} - I_{\text{Na,SOM}} - I_{\text{KDR,SOM}} - I_{\text{leak,SOM}} - I_{\text{exc}}$$

where $C_{\text{SOM}} = 0.9$ and $I_{\text{app,SOM}}$, the applied current, is constant in time. The form of each current is given in Table 1; equilibrium voltages are given in Table 2; and conductance values for all six models that will be introduced in *Results: Modeling cortical θ oscillators* (see Figure 1) are given in Table 3.

The dynamics of activation variable x (ranging over h, m_{KDR}, n, m_{Nap}, s, and q in Table 1) were given either in terms of its steady-state value x_∞ and time constant τ_x by the equation

$$\frac{dx}{dt} = \frac{x_\infty - x}{\tau_x},$$

Table 3. Maximal Conductances.

Model	M	MI	I	IS	MIS	MS
g_{Na}	125	125	125	125	125	125
$g_{\text{K}_{\text{DR}}}$	54	54	54	54	54	54
g_{leak}	0.31	0.27	0.78	0.78	0.27	0.27
g_{m}	1.4472	1.4472	0	0	1.4472	1.4472
g_{Na_P}	0.4307	0.4307	0.4307	0.4307	0.4307	0.4307
g_{Ca}	0.54	0.54	0.54	0.54	0.54	0.54
$g_{\text{K}_{\text{SS}}}$	0	0	0	0.1512	0.1512	0.1512
g_{app}	-7.1	-6.5	-7.6	-10.5	-9.8	-9.2
$g_{\text{Na,SOM}}$	0	100	100	100	100	0
$g_{\text{K}_{\text{DR}},\text{SOM}}$	0	80	80	80	80	0
$g_{\text{leak,SOM}}$	0	0.1	0.1	0.1	0.1	0
$I_{\text{app,SOM}}$	0	0.95	0.95	0.95	0.95	0
$g_{\text{RS}\to\text{SOM}}$	0	0.075	0.075	0.075	0.075	0
$g_{\text{SOM}\to\text{RS}}$	0	0.15	0.15	0.15	0.15	0

Table 4. Activation variable dynamics.

h	$\alpha_h(V) = 0.07\exp\left(-(V+30)/20\right)$	$\beta_h(V) = \left(\exp\left(-V/10\right)+1\right)^{-1}$
m_{Na}	$\alpha_m(V) = -\frac{V+16}{10(\exp(-(V+16)/10)-1)}$	$\beta_m(V) = 4\exp\left(-(V+41)/18\right)$
$m_{\text{K}_{\text{DR}}}$	$\alpha_m(V) = -0.01\frac{V+20}{\exp(-(V+20)/10)-1}$	$\beta_m(V) = 0.125\exp\left(-(V+30)/80\right)$
n	$n_\infty(V) = \left[1+\exp\left(-(V+35)/10\right)\right]^{-1}$	$\tau_n(V) = \frac{1000/(3.3*3^{(34-22)/10})}{\exp\left(\frac{V+35}{40}\right)+\exp\left(\frac{-(V+35)}{20}\right)}$
m_{Na_P}	$m_\infty(V) = \left[1+\exp\left(-(V+40)/5\right)\right]^{-1}$	$\tau_m = 5$
s	$\alpha_s(V) = 1.6\left(1+\exp\left(-0.072(V+65)\right)\right)$	$\beta_s(V) = 0.02\frac{V+51.1}{\exp\left(\frac{V+51.1}{5}\right)-1}$
q	$\alpha_q(C_{\text{Ca}}) = \min\left(0.1C_{\text{Ca}}, \quad 1\right)$	$\beta_q = 0.002$
h_{SOM}	$h_\infty(V) = \left[1+\exp\left((V+58.3)/6.7\right)\right]^{-1}$	$\tau_h(V) = 0.225 + 1.125\left[1+\exp\left((V+37)/15\right)\right]^{-1}$
$m_{\text{K}_{\text{DR}},\text{SOM}}$	$m_\infty(V) = \left[1+\exp\left((-V-27)/11.5\right)\right]^{-1}$	$\tau_m(V) = 0.25 + 4.35\left[1+\exp\left(-\lvert V+10\rvert/10\right)\right]^{-1}.$

or in terms of its forward and backward rate functions, α_x and β_x, by the equation

$$\frac{dx}{dt} = (1-x)\alpha_x - x\beta_x.$$

Only the expressions for m_{Na} differed slightly:

$$m_{\text{Na}}(V) = \alpha_m/\left(\alpha_m+\beta_m\right), \qquad m_{\text{Na,SOM}}(V) = \left[1+\exp\left((-V-38)/10\right)\right]^{-1}.$$

Steady-state values, time constants, and forward and backward rate functions are given in Table 4. For numerical stability, the backwards and forwards rate constants for q and s were converted to steady-state values and time constants before integration, using the equations

$$x_\infty = \alpha_x\tau_x, \qquad \tau_x = \left(\alpha_x+\beta_x\right)^{-1}.$$

The dynamics of the synaptic activation variable s were given by the equation

$$\frac{ds}{dt} = -\frac{s}{\tau_{\text{D}}} + \frac{1-s}{\tau_{\text{R}}}\left(1+\tanh\left(\frac{V_{\text{pre}}}{10}\right)\right),$$

with time constants $\tau_{\text{R}} = 0.25$ ms, $\tau_{\text{D,RS}\to\text{FS}} = 2.5$ ms, and $\tau_{\text{D,FS}\to\text{RS}} = 50$ ms.

F-I curves

For these curves, we varied the level of tonic applied current I_{app} over the range from 0 to 200 Hz, in steps of 1 Hz. We measured the spiking rate for the last 5 seconds of a 6 second simulation, omitting the transient response in the first second.

Table 5. Varied pulse input (I_{VP}) parameters (see *Methods: Phase-locking to rhythmic and quasi-rhythmic inputs: Inputs* for details).

Input Bandwidth ($= f_{\text{high}} - f_{\text{low}}$)	f_{low}	f_{high}	d_{low}	d_{high}	s_{low}	s_{high}	o_{low}	o_{high}
1	6.5	7.5	0.25	0.3	10	40	0	0.05
1.65	6.175	7.825	0.2375	0.325	10	41	0	0.1
2.3	5.85	8.15	0.225	0.35	9	41	0	0.15
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
13.35	0.325	13.675	0.0125	0.775	1	50	0	1

Phase-locking to rhythmic and quasi-rhythmic inputs

Inputs

In addition to the tonic applied current I_{app}, to measure phase-locking to rhythmic and quasi-rhythmic inputs, we introduced time-varying applied currents. These consisted either of periodic pulses (I_{PP}) or of variable-duration pulse trains with varied inter-pulse intervals (I_{VP}). Periodic pulse inputs were given by the expression

$$I_{\text{PP}}(t) = g_{\text{PP}}\Sigma_i \chi_{\left\{|t-t_i^*|<=w(s-1)/2s\right\}}(t) * \exp\left(-(st/w)^2\right), \tag{1}$$

where $\chi_S(t)$ is the function that is 1 on set S and 0 otherwise, $t_i^* = 2\pi\omega i$ for $i = 1, 2, \ldots$ is the set of times at which pulses occur, ω is the frequency, $w = 1000d/\omega$ is the pulse width given the duty cycle $d \in (0, 1)$, $*$ is the convolution operator, and s determines how square the pulse is, with $s = 1$ being roughly normal and higher s being more square. For our simulations, we took $d = 1/4$ and $s = 25$.

Variable-duration pulse trains were given by the expression

$$I_{\text{VP}}(t) = g_{\text{VP}}\Sigma_i \chi_{\left\{|t-t_i^*-o_i|<=w_i\frac{(s_i-1)}{2s_i}\right\}}(t) * \exp\left(-\left(\frac{s_i t}{w_i}\right)^2\right), \tag{2}$$

where

$$t_i^* = \Sigma_{j=1}^{i} 1000/\omega_j,$$

the frequencies $\{\omega_i\}_1^n$ are chosen uniformly from $[f_{\text{low}}, f_{\text{high}}]$, the pulse width is given by $w_i = 1000d_i/\omega_i$, the duty cycles $\{d_i\}_1^n$ are chosen uniformly from $[d_{\text{low}}, d_{\text{high}}]$, the shape parameters $\{s_i\}_1^n$ are chosen uniformly from $[s_{\text{low}}, s_{\text{high}}]$, and the offsets $\{o_i\}_1^n$ are chosen uniformly from $[o_{\text{low}}, o_{\text{high}}]$. For our simulations, these parameters are given in Table 5.

Phase-locking Value

The (spike rate adjusted) phase-locking value (PLV, **Aydore et al.** (**2013**)) of the oscillator to these inputs was calculated with the expressions

$$PLV = \left(n_s|MRV|^2 - 1\right) / \left(n_s - 1\right), \qquad MRV = \frac{1}{n_s}\sum_{i=1}^{n_s} \exp\left(\sqrt{-1}\phi_I\left(t_i^s\right)\right),$$

where MRV stands for mean resultant vector, n_s is the number of spikes, t_i^s is the time of the i^{th} spike, and $\phi_I(t)$ is the instantaneous phase of input (I_{PP} or I_{VP}) at frequency ω.

For I_{PP}, $\phi_I(t)$ was obtained as the angle of the complex time series resulting from the convolution of I_{PP} with a complex Morlet wavelet having the same frequency as the input and a length of 7 cycles. Since I_{VP} was composed of pulses and interpulse periods of varying duration, this procedure did not yield accurate estimates of the instantaneous phase of these inputs. Instead, the following procedure was used. First, the times that χ_{VP} went from zero to greater than zero $\left(\{a_i\}_{i=1}^n\right)$ and from greater than zero to zero $\left(\{b_i\}_{i=1}^n\right)$ were obtained. Second, we specified the phase of I_{VP} on these points via the function $\phi_I^0(t)$, a piecewise constant function satisfying

$$\frac{d}{dt}\phi_I^0(t) = \sum_{i=1}^{n}\left(\frac{3\pi}{2}\delta_{a_i}(t) + \frac{\pi}{2}\delta_{b_i}(t)\right),$$

where δ is the Dirac delta function. Finally, we determined $\phi_I(t)$ from $\phi_I^0(t)$ via linear interpolation, i.e. by setting $\phi_I(t)$ to be the piecewise linear (strictly increasing) function satisfying

$$\phi_I(0) = 0, \qquad \phi_I\left(a_i\right) = \phi_I^0\left(a_i\right), \qquad \phi_I\left(b_i\right) = \phi_I^0\left(b_i\right).$$

The resulting function $\phi_I(t)$ advances by $\pi/2$ over the support of each input pulse (the support is the interval of time over which the input pulse is nonzero), and advances by $3\pi/2$ over the time interval between the supports of consecutive pulses.

Spike-triggered input pulses

To explore the buildup of outward current and delay of subsequent spiking induced by strong forcing, we probed each model with a single spike-triggered pulse. These pulses were triggered by the first spike after a transient interval of 2000 ms, had a pulse duration of 50 ms, and had a form given by the summand in Equation 1 with $w = 50$ and $s = 25$ (i was 1 and t_i was the time of the triggering spike).

Results

Modeling cortical θ oscillators

To explore how frequency flexibility in phase-locking depends on the biophysics and dynamics of inhibitory currents, we employed Hodgkin-Huxley type computational models of cortical θ oscillators. In these models, θ rhythmicity was paced by synaptic inhibition with a fast rise time and a slow decay time and/or by θ-frequency sub-threshold oscillations (STOs) resulting from the interaction of a pair of intrinsic currents activated at subthreshold membrane potentials – a depolarizing persistent sodium current and a hyperpolarizing and slowly activating m-current (***Gutfreund et al., 1995***) (Fig. 1, 2B). A super-slow potassium current introduced a δ timescale into the dynamics of some models (Fig. 1, 2C). Thus, in addition to spiking and leak currents, our models included up to three types of outward – i.e. hyperpolarizing and thus spike suppressing, and here termed inhibitory – currents: an m-current or slow potassium current (I_{m}) with a voltage-dependent time constant of activation of ∼10-45 ms; recurrent synaptic inhibition (I_{inh}) with a decay time of 60 ms; and a super-slow K current (I_{KSS}) with (calcium-dependent) rise and decay times of ∼100 and ∼500 ms, respectively. The presence of these three hyperpolarizing currents was varied over six models – M, I, MI, MS, IS, and MIS – whose names indicate the presence or absence of each current: M for the m-current, I for synaptic inhibition, and S for the super-slow K current (Fig. 1).

To parameterize our models, we began by qualitatively matching *in vitro* recordings from layer 5 θ-resonant pyramidal cells (***Carracedo et al., 2013***) (Fig. 2D). These RS cells transition from tonic δ-rhythmic spiking to tonic θ-rhythmic spiking through so-called mixed-mode oscillations (MMOs, here pairs of spikes spaced a θ period apart occurring at a δ frequency) as their resting membrane potential is raised over a few mV (***Carracedo et al., 2013***), and the *in vitro* data suggests that this pattern of spiking is independent of recurrent synaptic inhibition, arising instead from intrinsic inhibitory currents. To replicate this behavior, we constructed a Hodgkin-Huxley neuron model paced by both I_{m} and I_{KSS} (Fig. 1, model MS, & Fig. 2A). While *in vitro*, these cells receive δ-rhythmic EPSPs, this rhythmic excitation is not required in our model, which exhibited MMOs in response to tonic input (Fig. 2D).

We then constructed five additional models based on model MS (Fig. 1). First, to obtain model IS, we replaced I_{m} with I_{inh}, adjusting the leak current and the conductance of synaptic inhibition to get a frequency-current (FI) curve having a rheobase and inflection point similar to that of model MS. In the remaining models, only the leak current conductance was changed to produce θ-rhythmic spiking at roughly similar values of I_{app}; all other conductances were identical to those in models MS and IS. For all remaining simulations, we fixed I_{app} so that all models exhibit spontaneous rhythmic spiking at ∼7 Hz (Fig. 1, red circles).

Phase-Locking Under Strong Forcing

We first tested whether these six oscillators entrained to different frequencies of exactly periodic input, examining their responses to rhythmic input pulses with frequencies ranging from 0.25 to 23 Hz. To mimic the bursts of excitation produced by deep intrinsic bursting (IB) cells projecting to deep regular spiking (RS) cells (***Carracedo et al., 2013***), rhythmic inputs were modeled as smoothed square-wave current injections to the RS cells of all three models, having duty cycles 1/4 of the input period (see Methods). We varied the strength of input pulses from 0 to 4 pA, and measured the degree of phase-locking to the input rhythm exhibited by RS cell spikes over 30 seconds (see Methods). The results of these simulations are shown in Fig. 3, with models ordered by increasing frequency flexibility of phase-locking. For high enough input strength, all models were able to phase-lock adequately to inputs faster than 7 Hz, including the fastest frequency we tested (23 Hz).

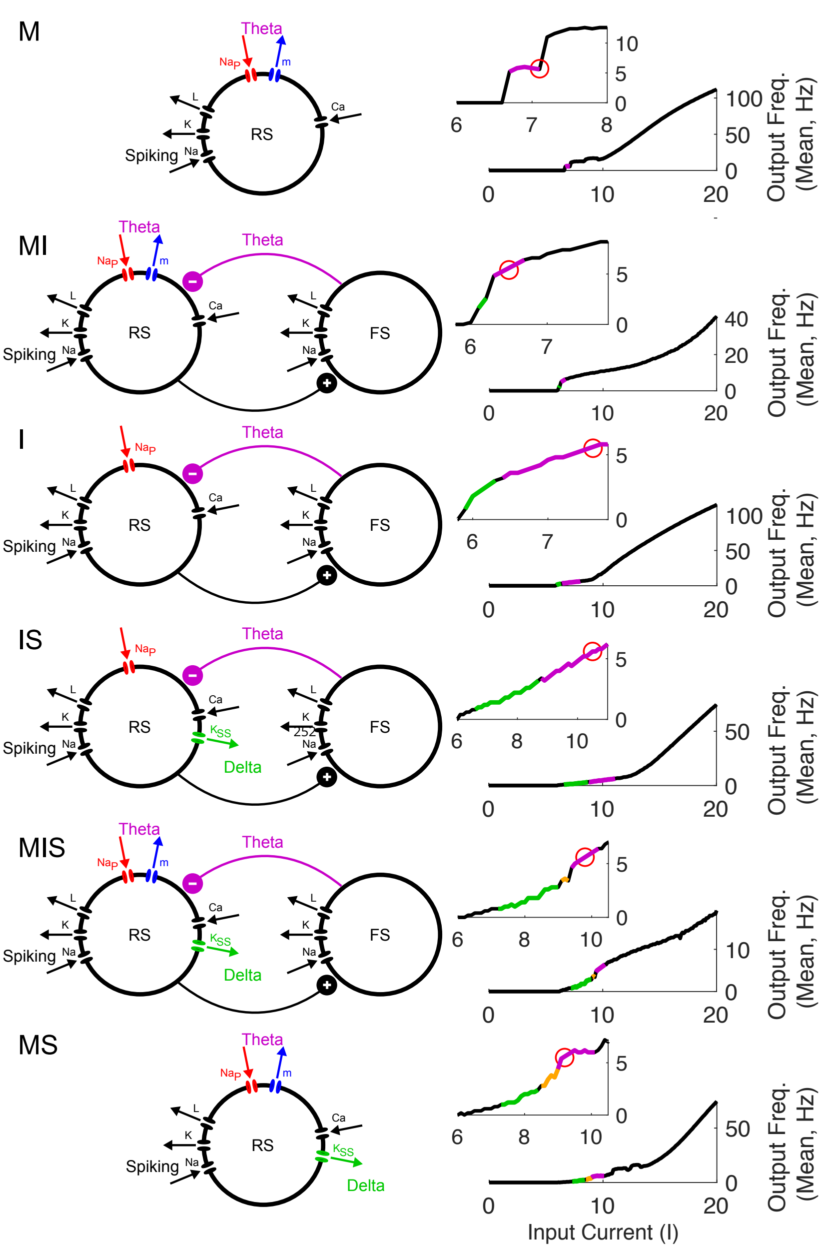

Figure 1. Model θ oscillators. Schematics (left) show the currents present in each model, color-coded according to timescale, with δ in green and θ in purple. FI curves (right) show the transition of spiking rhythmicity through δ and θ frequencies as I_{app} increases, with δ in green, θ in purple, and δ-nested θ (i.e., mixed-mode oscillations) in gold; the red circle indicates the applied current giving a firing rate of ∼7 Hz.

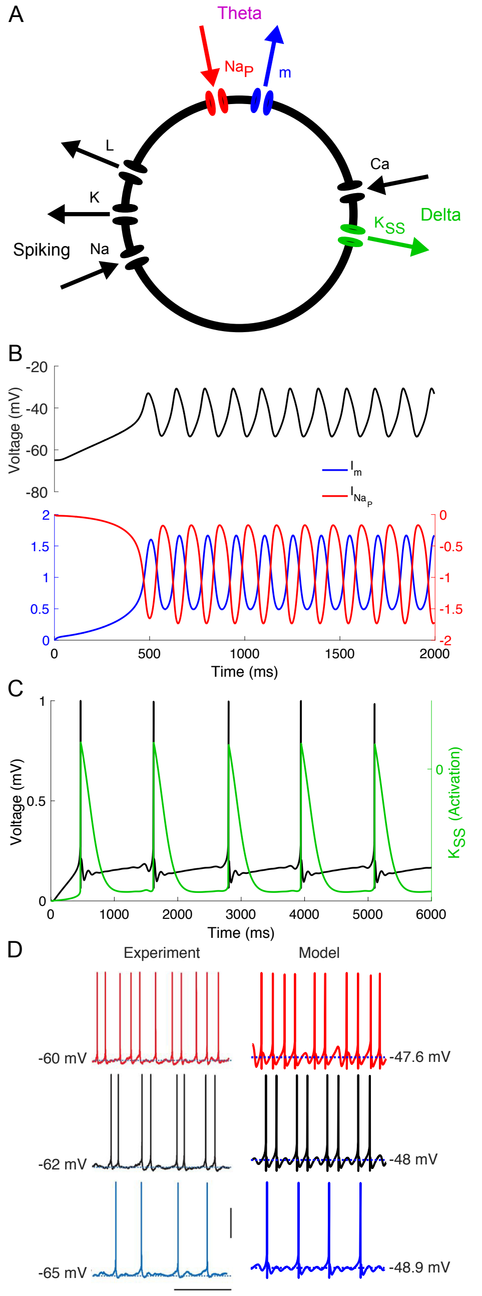

Figure 2. Model MS reproduces *in vitro* data. (A) Diagram of model MS. Arrows indicate directions of currents (i.e., inward or outward). (B) θ timescale STOs arise from interactions between m- and persistent sodium currents in a model without spiking or Ca-dependent currents (only g_{m} and g_{Na_P} nonzero). (C) δ timescale activity-dependent hyperpolarization arises from a calcium-activated potassium current. (D) Comparison between *in vitro* (adapted from (***Carracedo et al., 2013***)) and model data (calibration: 50 μV, 0.5 ms).

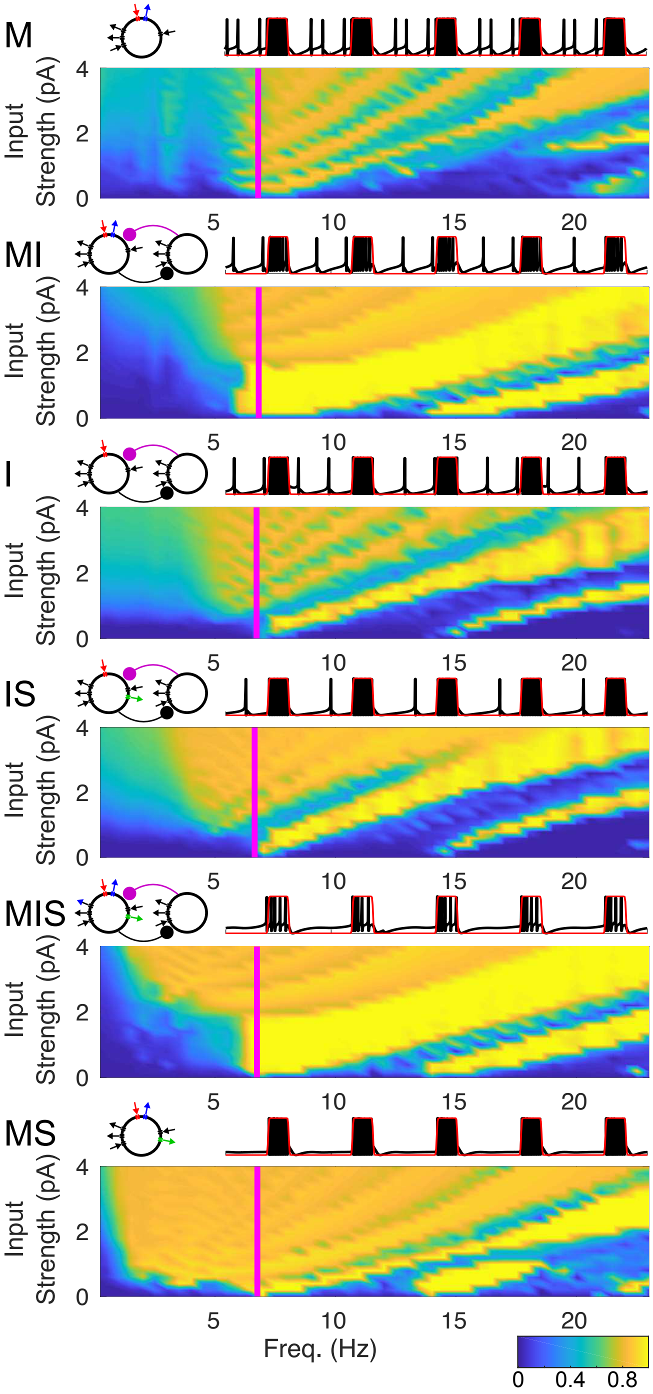

Figure 3. Phase-Locking as a Function of Periodic Input Frequency & Strength. False-color images show the (spike-rate adjusted) phase-locking value (PLV, see Methods) of spiking to input waveform, with vertical magenta lines indicating intrinsic spiking frequency. Schematics of each model appear above and to the left; sample traces of each model appear above and to the right (voltage traces in black, input profile in red, two seconds shown, input frequency 2.5 Hz, input strength -3.4 pA). The bands in these false-color images are related to the number of spikes generated per input cycle: the highest PLV occurs when an oscillator produces one spike per input cycle, and PLV decreases (from band to band) as the strength of the input (and the number of spikes per input cycle) increases.

Models exhibited significantly more diversity in their ability to phase-lock to inputs slower than their intrinsic frequency, ranging from total inability to phase-lock to input frequencies below 7 Hz

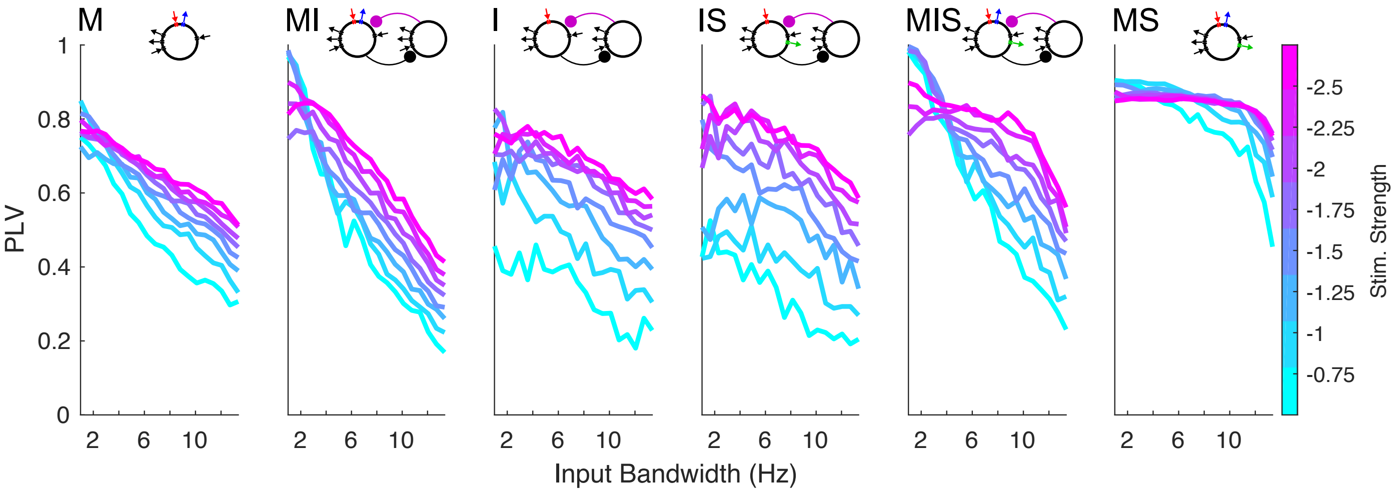

Figure 4. Phase-Locking to Quasi-Rhythmic Inputs. Plots show the (spike-rate adjusted) phase-locking value of spiking to input waveform, for inputs varying in bandwidth and regularity having a center frequency of 7 Hz (see Methods & Table 5) as well as input strength. Schematics of each model appear above and to the left.

(model M), to ability to phase lock to input frequencies as low as 1.5 Hz even when input strength was relatively low (model MS). The super-slow K current uniformly increased the frequency range of phase-locking, with every model containing I_{KSS} able to phase-lock to slower inputs than any model without I_{KSS}. Synaptic inhibition seemed to stabilize the frequency range of phase-locking, with the four models containing I_{inh} exhibiting an intermediate frequency range of phase-locking, while the m-current seemed to do the opposite, with both the narrowest and the broadest frequency ranges of phase-locking occurring in the four model θ oscillators containing I_{m}, and the very narrowest and broadest ranges occurring in the two of these models that lacked I_{inh} (models M and MI). Notably, models MI and MIS exhibited one-to-one phase-locking to periodic inputs (i.e., a single spike per input cycle occurring within a small window of input phases, appearing in Fig. 3 as a bright yellow band indicative of high PLV) for input strengths twice as high as any other models. In these models, synaptic inhibition and m-current inhibition were both present, making the overall level of inhibition higher, and simulations showed that these phenomena were related: models exhibited one-to-one phase-locking increased over larger ranges of input strengths as the conductances of the m- and synaptic currents were increased (Fig. S1).

Next, we tested whether the frequency selectivity of phase-locking exhibited for periodic inputs would carry over to quasi-rhythmic inputs, by exploring the phase-locking behavior of model θ oscillators in response to trains of input pulses in which pulse duration, interpulse duration, and pulse waveform varied from pulse to pulse (with a center frequency of 7 Hz). These irregular input pulse trains probed the abilities of our θ oscillator models to "parse" irregular inputs. To create these irregular trains of input pulses, pulse and interpulse durations were chosen (uniformly) randomly from ranges of pulse "frequencies" and "duty cycles", respectively, and pulse shape and onset time were similarly randomized (see Methods, Equation 2). To create a gradient of input classes with different degrees of regularity, we systematically varied the intervals from which frequencies (determining cycle lengths), duty cycles, pulse shapes, and pulse onset times were chosen (Table 5); we use "bandwidth" here as a shorthand for this multi-dimensional gradient in input regularity.

For "narrowband" (i.e. highly regular) inputs and high input strengths, all six models showed a high degree of phase-locking to the input waveforms, reflecting their ability to phase-lock to periodic inputs at their intrinsic frequency (Fig. 4). In contrast, phase-locking to "broadband" inputs (i.e., irregular inputs with broad ranges of input periods, durations, and shapes) was higher for the models that exhibited broader frequency ranges of phase-locking to strictly periodic inputs – namely, models exhibiting both I_{KSS} and I_{m} (i.e., MIS and MS). At high input strengths, model MS in particular showed a high level of phase-locking that was nearly independent of input regularity (Fig. 4). Notably, model MIS mirrored the ability of model MS to phase-lock to broadband inputs at high input intensity, while showing frequency selective phase-locking at low input intensity. Indeed, MIS phase-locked to weak, narrowband quasi-rhythmic inputs better than any other model, perhaps due to its large region of one-to-one phase-locking (Fig. 4).

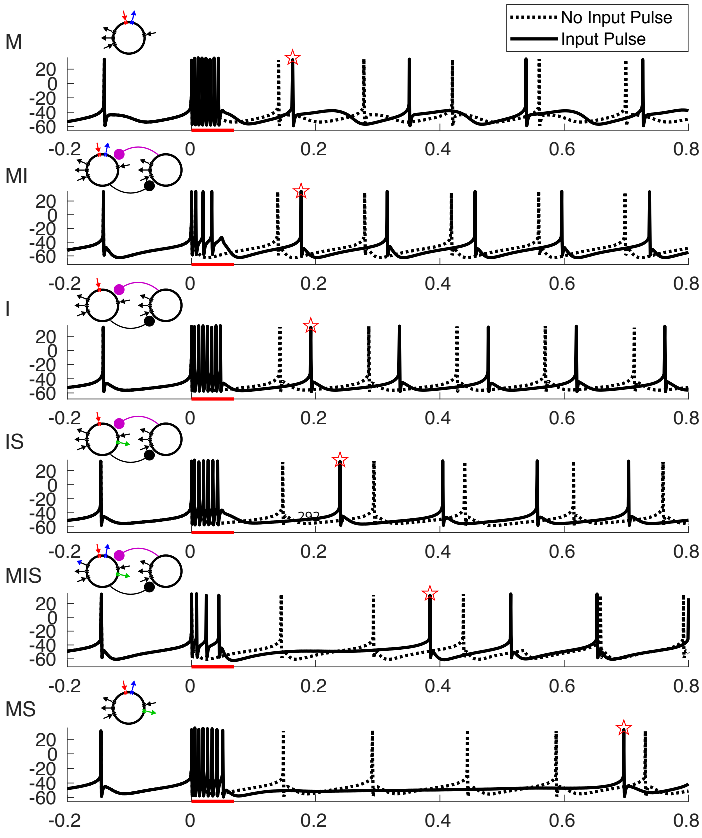

Figure 5. Delay of spiking in response to input pulses. Voltage traces plotted for simulations both with (dotted lines) and without (solid lines) an input pulse lasting 50 ms. Red bar indicates the timing of the input pulse; red star indicates the first post-input spike.

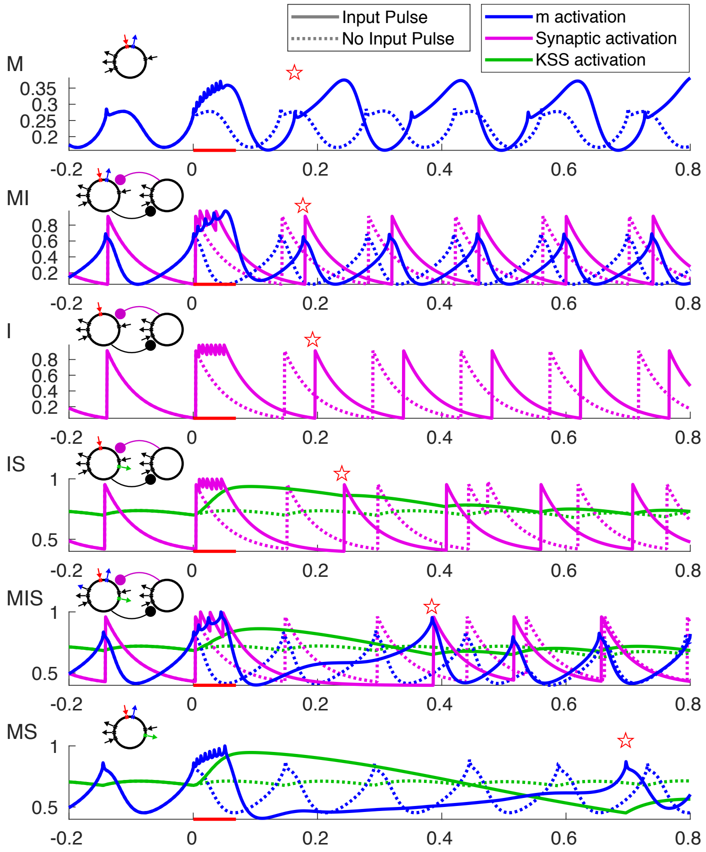

Figure 6. Buildup of outward currents in response to input pulses. Activation variables (color) plotted for simulations both with (dotted lines) and without (solid lines) an input pulse lasting 50 ms. Red bar indicates the timing of the input pulse; red star indicates the time of the first post-input spike.

Buildup of Outward Currents

Next, we sought to understand how the dynamics of inhibitory currents contributed to the observed gradient from selective to flexible phase-locking, and in particular to the observed variations in entrainment to rhythms with frequencies slower than θ. To begin answering this question, we discuss the mechanism of phase-locking to periodic input pulse trains in our models. During each input pulse, strong forcing leads to a burst of spiking, which in turn activates the outward currents that pace the models' intrinsic rhythmicity. These inhibitory currents hyperpolarize the models, causing the cessation of spiking for at least a θ period, and in some cases much longer. If the pause in spiking is sufficiently long to delay further spiking until the next input arrives, phase-locking is achieved, given that the next strong input pulse causes spiking. Thus, the strength and duration of the post-input hyperpolarization determines the duration of the pause in spiking, and this in turn determines the lower (frequency) limit of phase-locking to periodic inputs for the oscillator, which, as we have shown, is related to the frequency flexibility of phase-locking to irregular inputs.

The dynamics of intrinsic and synaptic currents determine the length of this pause and its dependence on input strength. To observe how the dynamics of outward currents contributed to the length of the delay in spiking following an input pulse, we explored model responses to single (excitatory) input pulses (Fig. 5). As expected, the delay of spiking in response to a single, strong input pulse lasting 50 ms corresponded overall to the frequency flexibility of phase-locking in our models, being shortest in model M and longest in model MS (Fig. 5). The gating variables of the three outward currents simulated in our models offered an explanation for the observed patterns of phase-locking frequency flexibility (Fig. 6). We describe the dynamics of these currents during and after the input pulse, from simple to complex.

The simplest current to understand is I_{inh} (Fig. 6, purple), which does not build up during the input pulse, and by itself (in model I) delays the spike following the input pulse by very little. This small delay occurs because the input pulse, by causing the RS and then the SOM cell to spike, repeatedly resets the synaptic inhibitory "clock" - the time until I_{inh} has decayed enough for a spontaneous spike to occur. As soon as spiking stops (at the end of the input pulse or shortly afterwards - our model SOM interneurons are highly excitable and may exhibit noise-induced spiking after the input pulse), the level of inhibition begins to decay.

As seen in model M (Fig. 6), the m-current (blue) does build up over the course of the input pulse, but it quickly decays at the end of the input pulse, as m-current activation continues to oscillate. As a result, the spike that follows the input pulse is delayed very little, relative to the spiking behavior of the unforced oscillator (Fig. 5). Thus, the STOs in model M offer an explanation for this model's inability to phase-lock to rhythms below its intrinsic frequency. While the buildup of the m-current during the input pulse seems to change the frequency of the subsequent STOs, this doesn't affect the model's phase-locking in the strong forcing regime, which depends primarily on the length of the pause in spiking following the input pulse.

The super-slow K current builds up dramatically during the input pulse (Fig. 6, green), and decays slowly, increasing the latency of the first spike following the input pulse substantially (Fig. 5). This slow-building outward current interacts differently, however, with synaptic and intrinsic θ-timescale currents.

In model IS, both I_{inh} and $I_{\text{K}_{\text{SS}}}$ decay monotonically following an input pulse, until the total level of hyperpolarization is low enough to permit another spike. The spike burst in response to the input pulse means that the effective level of inhibition due to I_{inh} is much higher than the effective level of inhibition due to $I_{\text{K}_{\text{SS}}}$; thus, the delay in spiking following the input terminates after the synaptic activation variable s has decayed below the level $s*$ at which the cell spikes spontaneously, but well before the activation of $I_{\text{K}_{\text{SS}}}$ has returned to spontaneous levels (Fig. 6, model IS). Nevertheless, $I_{\text{K}_{\text{SS}}}$ and I_{inh} appear to interact additively to produce hyperpolarization and a pause in spiking in the RS cell.

In model MS, the conductance of I_{m} is again much higher than the conductance of $I_{\text{K}_{\text{SS}}}$. However, in this model, the $I_{\text{K}_{\text{SS}}}$ activation must decay to levels much lower than "baseline" before the oscillator spikes again, even though there are times during the delay in spiking following the input pulse at which both I_{m} and $I_{\text{K}_{\text{SS}}}$ are less active than during baseline spiking (e.g., ~0.45 and ~0.6 seconds, Fig. 6). Notably, the m-current activation increases for the duration of the delay period, with subtle oscillatory fluctuations riding the increase. These dynamics suggest a more complex relationship between I_{m} and $I_{\text{K}_{\text{SS}}}$. While a full analysis is outside the scope of this paper, we hypothesize that this synergistic effect is due to the STOs present in model MS.

In model MI, buildup of I_{m} decreases SOM cell spiking during the input pulse, interfering with the "resetting" of I_{inh} near the end of the input pulse, while the presence of I_{inh} damps the STOs in the RS cell, resulting in behavior intermediate between models M and I. Similarly, model MIS exhibits behavior intermediate between models MS and IS. Following an input pulse, spiking is delayed until $I_{\text{K}_{\text{SS}}}$ decays below baseline levels (Fig. 6) - a longer delay than in model IS. It seems unlikely that the three outward currents interact additively in model MIS, since, as in model MS, the first post-input spike occurs at a time at which the activation level of the m-current is elevated relative to baseline (Fig. 6).

Discussion

Our results show that multiple timescales of synaptic and intrinsic inhibition interact in dynamically complex ways. In particular, we discovered a complex relationship between δ- and θ-timescale intrinsic K currents that enables some of our model θ oscillators to follow strong periodic forcing well below their intrinsic spiking frequency. Similar oscillators lacking this pair of intrinsic currents (including oscillators paced primarily by synaptic inhibition) show markedly less frequency flexibility in phase-locking to rhythmic and quasi-rhythmic inputs. Thus, synaptic and intrinsic inhibition may tune neural oscillators to exhibit different levels of phase-locking flexibility, allowing them to play diverse roles – from reliable timing cues to flexible parsers of sensory inputs.

Mechanisms of phase-locking

Our results show that, for models containing a variety of intrinsic and synaptic currents, buildup of outward current during strong forcing plays an important role in the ability to phase-lock to inputs having a wide range of frequencies, especially frequencies below the models' spontaneous frequency. A super-slow adapting current, $I_{\mathrm{K_{SS}}}$, is key to the ability to phase-lock to slow frequencies in our models. In general, whether an outward current functions as an adapting current – causing the time intervals between successive pairs of spikes to decrease during intense stimulation – depends on the timescale of the current relative to the timescale of spiking during stimulation. If inhibition or hyperpolarization is activity-dependent and has a timescale slower than the interspike interval, it can serve this function; for example, in a γ oscillator, any current with a timescale below γ (for example, the θ timescale m-current) could function in this way.

We have also shown, we believe for the first time, a synergy in our models between a slow and a super-slow K current. While a full analysis is beyond the scope of this paper, we conjecture that the synergy depends on the subthreshold oscillations (STOs) engendered by the slow K current (the m-current) in our models, as is suggested by a comparison between our IS, MIS, and MS models. In model IS, there are no STOs, and the interaction between θ-timescale inhibition (which is synaptic) and $I_{\mathrm{K_{SS}}}$ is additive. In models MIS and MS, where STOs result from interactions between the m-current and the persistent sodium current, the interaction between $I_{\mathrm{K_{SS}}}$ and θ-timescale hyperpolarization (mediated by the m-current) is not additive, but synergistic, producing a longer delay than would result from the sum of the two phenomena acting in isolation.

The fact that our oscillators can phase-lock to quasi-rhythmic as well as periodic inputs is largely a consequence of the mechanism of phase-locking in the regime of strong forcing. Since inputs are generally strong enough to cause spiking, phase-locking is dependent less on the phase at which an input arrives than on the delay caused by each input. If an input arrives before the end of that delay, it causes a spike, and phase-locking occurs. However, decreased phase-locking for inputs having a large range of instantaneous frequencies occurs in our models due to both “missed" spikes during the high-excitability phase of the input and “extra" spikes during the low-excitability phase of the input. These “missed" spikes may be due to the properties of our oscillators in the weak-forcing regime (see *Discussion: Relationship to previous work*).

Our model MIS is perhaps the most physiologically realistic, in that neurons in deep cortical layers are likely to exhibit all three outward currents studied in this paper. Significantly, this model exhibits both frequency selectivity in phase-locking at low input strengths, and frequency flexibility in phase-locking at high input strengths (Fig. 4). Input gain can depend on a variety of factors, including attention, stimulus novelty and salience, and whether the input is within- or cross-modality. A mechanism that allows input gain to determine the degree of phase-locking frequency flexibility could enable the differential processing of inputs based on these attributes.

Relationship to previous work

Phase-locking of neural oscillators under weak forcing has been studied extensively (***Ermentrout, 1981***, ***1996***; ***Kopell and Ermentrout, 2002***; ***Achuthan and Canavier, 2009***; ***Canavier and Achuthan, 2010***). In this regime, a neural oscillator stays close to a limit cycle during and after forcing, and as a result the phase of the oscillator is well-defined throughout forcing. Furthermore, the change in phase induced by an input is small (less than a full cycle), can be calculated, and can be plotted as a function of the phase at which the input is applied, resulting in a phase-response curve (PRC). In this work, we have focused on strong and non-instantaneous forcing. Our results pertain to a dynamical regime in which PRC theory does not apply. Namely, our forcing is strong and long enough that our oscillators complete multiple cycles during the input pulse, and as a result the phase at the end of forcing is not guaranteed to be a function of the phase at which forcing begins. Furthermore, in oscillators which contain $I_{\mathrm{K_{SS}}}$, the dynamics of this slow current adds an additional dimension, which makes it impossible to describe the state of these oscillators in terms of a simple phase variable. Not only the phase of the oscillator, but also its amplitude (which is impacted by the activation of $I_{\mathrm{K_{SS}}}$), determine its dynamics.

We have focused on the θ timescale, and our results highlight in particular the properties of the θ-timescale m-current. Previous work has illuminated many of the dynamical properties of the m-current. The addition of an m-current (or indeed any slow resonating current, such as an h-current or a slow, non-inactivating potassium current) changes a neuron from a Type I to a Type II oscillator (***Ermentrout et al., 2001***; ***Acker et al., 2003***). The fact that resonating currents can generate membrane potential resonance (and subthreshold oscillations) is well-studied (***Gutfreund et al., 1995***; ***Hu et al., 2002***; ***Rotstein and Nadim, 2014***). More recently, it has been shown that the θ-timescale properties of the M-current allow an E-I network subject to θ forcing to precisely coordinate with external forcing on a γ timescale (***Zhou et al., 2018***). While STOs play a role in the dynamics of our model oscillators containing the m-current, subthreshold resonance does not automatically imply suprathreshold resonance or precise response spiking (***Rotstein, 2017***). Thus, our results are not predictable (either *a priori* or *a posteriori*) from the effects of the m-current on neuronal dynamics.

It is possible that larger inhibition-paced networks, which have been studied both computationally and experimentally (***Akam and Kullmann, 2012***; ***Tsai et al., 2008***; ***Atallah and Scanziani, 2009***; ***Shin and Cho, 2013***; ***Sherfey et al., 2018a***), may exhibit properties distinct from our models containing synaptic inhibition. Papers attempting to spell out the constraints on selective communication between neuronal oscillators have examined larger networks (***Akam and Kullmann, 2012***; ***Sherfey et al., 2018a***). Other computational work has shown that the addition of E-E and I-I connectivity in E-I networks can yield frequency flexibility through potentiation of these recurrent connections (***Tsai et al., 2008***; ***Shin and Cho, 2013***), although the putative timescale of this potentiation is much slower than the cycle-by-cycle changes in instantaneous frequency seen in physiologically-relevant quasi-rhythmic inputs. Faster mechanisms may be at play in the brain, however: experimental results show that amplitude and instantaneous frequency are related in hippocampal networks, since firing by a larger proportion of excitatory pyramidal cells recruits a larger population of inhibitory interneurons (***Atallah and Scanziani, 2009***). In larger E-I networks, the recruitment of different proportions of neurons by inputs of different sizes may allow more frequency flexibility in phase-locking; and recent work shows something similar: that heterogeneity in E-I synapses can allow frequency flexibility in phase-locking recruiting different subpopulations of interneurons (). Finally, in our models, as in published models of θ-oscillator driven syllable segmentation (***Hyafil et al., 2015***), only the RS cells receive inputs. How feedforward inhibition affects the frequency specificity of phase-locking, especially in larger networks, is an important question for future research.

Functional implications

Our focus on the θ timescale is motivated by *in vitro* results on cortical θ oscillators (***Carracedo et al., 2013***), by *in vivo* results underscoring the prominence of theta rhythms in the spontaneous activity of primate auditory cortex (***Lakatos et al., 2005***), and by the potential functional role of cortical θ oscillations in segmentation auditory input (***Ghitza and Greenberg, 2009***; ***Ghitza, 2011***, ***2012***, ***2013***, ***2014***; ***Riecke et al., 2015b***,a, ***2017***; ***Teng et al., 2017***; ***Zoefel et al., 2018***). Auditory cortical θ rhythms modulate the stimulus-specific response gain of auditory population activity (***Kayser et al., 2015***), and effect a continuous and active segmentation and grouping process for auditory stimuli (***Teng et al., 2017***). The phase entrainment between brain rhythms and speech envelope fluctuations at δ/θ frequencies is associated with speech intelligibility (***Ahissar et al., 2001***; ***Luo and Poeppel, 2007***; ***Nourski et al., 2009***; ***Hertrich et al., 2012***; ***Peelle et al., 2012***; ***Doelling et al., 2014***; ***Ding et al., 2016***; ***Riecke et al., 2017***; ***Zoefel et al., 2018***) and alters auditory perception of non-speech stimuli (***Zoefel and VanRullen, 2015***). Recent experiments have shown a causal link (***Riecke et al., 2017***; ***Wilsch et al., 2017***, ***2018***; ***Zoefel et al., 2018***): application of ~3 Hz transcranial alternating current stimulation at different phases relative to a rhythmically-presented speech stream phasically modulates behavior and the BOLD response to intelligible speech (***Zoefel et al., 2018***).

In terms of speech comprehension, the information in syllabic rate amplitude fluctuations is critical: removing 1-7 Hz (***Elliott and Theunissen, 2009***) or 2-9 Hz (***Ghitza, 2012***) amplitude modulations decreases speech intelligibility dramatically (***Drullman et al., 1994***), and periodic speech manipulations including interruption (***Miller and Licklider, 1950***), alternating monaural sources (***Huggins, 1964***), and reversal of short segments (***Stilp et al., 2010***) affect comprehension most when interfering with syllabic-rate intervals. Perhaps the strongest evidence supporting syllabic segmentation by θ oscillations is the observation that the phase of intrinsic ~6 Hz oscillations at sound onset determines the categorization of ambiguous syllables (between /da/ and /ga/), by altering their perceived onset length (***Ten Oever and Sack, 2015***).

The average spoken syllable lasts about the period of a 3 Hz oscillation, and temporal compression increasing the syllabic rate above 9 Hz results in a sharp drop in speech intelligibility that can be rescued by “repackaging" – inserting gaps of silence into the speech signal – with the highest levels of comprehension occurring when 333 ms segments of natural speech are delivered at rates below 9 Hz (***Ghitza and Greenberg, 2009***). Recent research shows that cortical speech-brain entrainment occurs for syllabic rates as high as 13 Hz, a speed at which speech is unintelligible, while β-frequency activity is abnormal in response to this unintelligible compressed speech (***Pefkou et al., 2017***). This suggests that the upper syllabic rate limit on speech intelligibility arises from the timescale of mnemonic processing (with a β-rhythmic signature) in structures downstream from the cortical θ oscillators responsible for syllabic segmentation (***Pefkou et al., 2017***). This is in agreement with our finding that the upper frequency boundary on phase-locking for our models extends well above 9 Hz, and is largely determined by input strength. Nonetheless, it is noteworthy that task-related auditory cortical entrainment operates most reliably over the 1-9 Hz (syllabic) ranges (***Lakatos et al., 2013***).

For the accurate segmentation of syllables, we are not interested in the ability of an oscillator to phase-lock, *per se*, to an input quasi-rhythm in the θ syllable range. Rather, what is required is a system that can reliably mark *every* syllable boundary – where syllable boundaries tend to occur quasi-rhythmically. Thus, we want a cellular oscillator which can reliably generate a single spike or a spike volley, at every high-energy phase that occurs within a θ quasi-rhythm. Our model MS seems capable of performing this task, if sufficiently excited, although it will be important to test this with actual speech waveforms. However, the phase-locking properties of model MS depend on the relationship between input strength and tonic excitation, and there is a tradeoff: at lower levels of tonic excitation, a single spike marks the peak of each low-frequency cycle, but some peaks of high-frequency inputs go unmarked, while at higher levels of tonic excitation, single spikes mark the peaks of high-frequency inputs, but multiple spikes may occur during the peaks of each low-frequency cycle (Fig. S2). Thus, it seems possible that the brain relies on a *bank* of such oscillators with different resting membrane potentials and/or input gains, connected synaptically or electrically, to more reliable mark syllable transitions across the entire “θ-syllable" range. As noted above, larger E-I networks may also exhibit more frequency flexibility than the inhibition-paced oscillators modeled here. The full spectrum of relationships between the properties of these individual oscillatory units and the activity of the networks in which they participate remains to be mapped out. These questions are complicated by the possibility that neuromodulators which affect the conductance of outward currents could change the flexibility vs. selectivity of phase-locking in individual oscillatory units; acetylcholine is one neuromodulator known to alter the conductance of the m-current. Nevertheless, we may ask why the brain would select phase-locking flexibility in single cells vs. networks. One possible answer is energetic efficiency. If flexibility in an inhibition-paced oscillatory network depends on recruiting large numbers of inhibitory interneurons, it may be more efficient as well as more reliable to utilize a small number of oscillators, each capable of segmentation quasi-rhythmic inputs containing a large range of instantaneous frequencies.

Entrainment flexibility does not seem to be ubiquitous in the brain. In the hippocampus, for example, θ rhythm is robustly periodic, exhibiting relatively small frequency changes with navigation speed (***McFarland et al., 1975***). Interestingly, the mechanisms of hippocampal θ and the neocortical θ rhythmicity discussed in this paper are very different: while the former is dominated by synaptic inhibition, resulting from an interaction of synaptic inhibition and the h-current in oriens lacunosum moleculare interneurons (***Rotstein et al., 2005***), the latter is only modified by it (***Carracedo et al., 2013***). Our results suggest that mechanisms like that of hippocampal θ, far too inflexible to perform the segmentation tasks necessary for speech comprehension, are instead optimized for a different functional role. One possibility is that imposing a more rigid temporal structure on population activity may help to sort “signal" from “noise" - i.e., imposing a strict frequency and phase criterion that inputs must meet to be processed. Another possibility is that more rigidly patterned oscillations result from a tight relationship to motor sampling routines which operate over an inherently more constrained frequency range, as, for example, whisking, sniffing, and running are related to hippocampal θ (***Kleinfeld et al., 2006***, ***2016***).

The auditory cortical θ oscillator is hypothesized to drive γ circuits, and in a phenomenological model matching psychophysical data (***Ghitza, 2011***), the frequency of γ is required to scale with the frequency of θ. Meeting this requirement strictly is acausal - the γ oscillator would have to access the duration of a θ cycle at the start of that cycle, before the duration of the cycle is, in fact, determined. However, we suspect it is possible for this requirement to be met with some lag, if the θ oscillator excites a γ oscillation via synaptic or intrinsic currents having a θ timescale. In this case, excitation would build up over multiple θ cycles when syllables arrive at a frequency faster than the timescale of excitation, and it would drop when syllables arrive at a frequency slower than the timescale of excitation. By similar mechanisms, slower oscillations, such as δ oscillations, could affect the intrinsic frequency of our θ oscillator. The current work is part of a larger research plan to develop a computational model of the oscillatory hierarchy that contributes to speech segmentation via phase-locking of brain rhythms to speech stimulus features. We hope this work will shed light on the biophysical mechanisms used by the brain to parse a complex auditory stimulus evolving on multiple timescales, namely, human speech.

Acknowledgments

BRPP, DAS, CES, and NJK were supported by NIH P50-MH109429. MAW was supported by Wellcome Trust #098353. CES was supported by NIH R01-MH111439. We thank Oded Ghitza for many useful discussions.

Supplemental Material

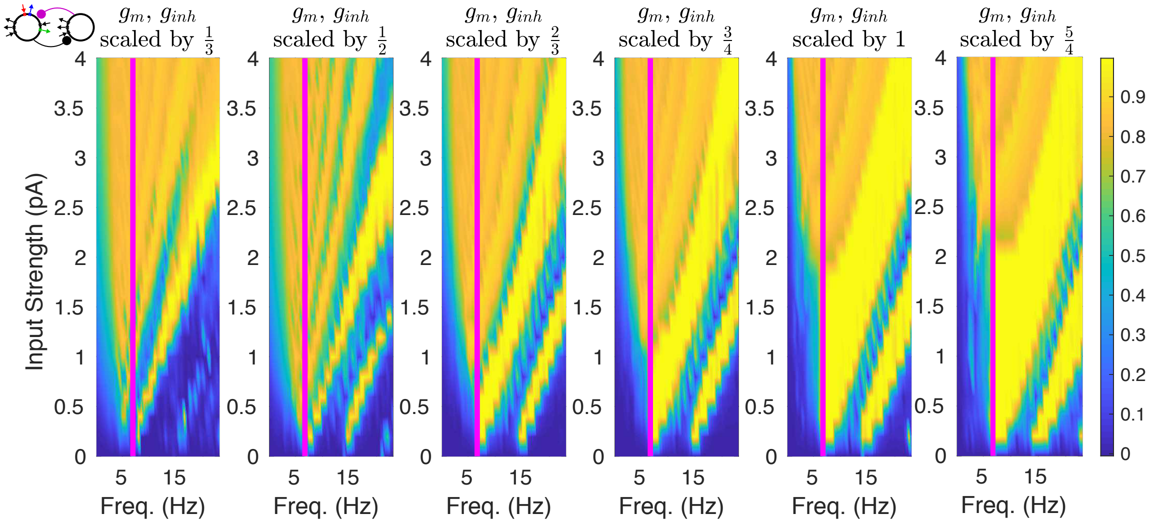

Figure S1. Dependence of One-to-One Phase Locking on Inhibitory Conductance. We multiplied the conductances g_m and g_inh in model MIS by factors of $\frac{1}{3}$, $\frac{1}{2}$, $\frac{3}{4}$, 1, and $\frac{5}{4}$, and then computed plots of PLV for different input frequencies and strengths, as in Fig. 3. The bright yellow band in each figure, representing the region of one-to-one phase-locking, depends on the size of g_m and g_inh; both increase from left to right.

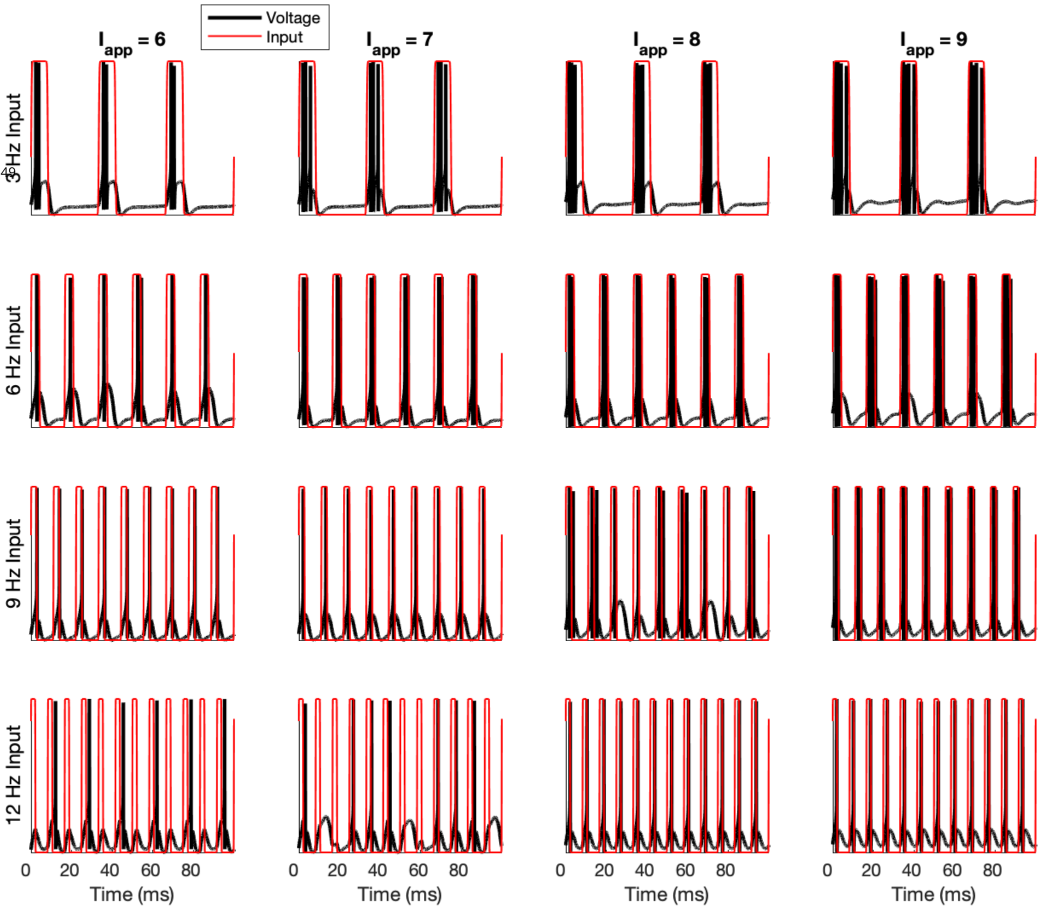

Figure S2. Varying Tonic Input to Model MS. We altered the tonic input strength g_app to model MS, and gave periodic pulse inputs of strength $g_\mathrm{PP} = 1$ at varying frequencies. For lower levels of tonic input, phase-locking is closer to one-to-one for low frequency inputs, but many high frequency input cycles are “missed”; for higher levels of tonic input, phase-locking is one-to-one for high frequency inputs, but many-to-one for low frequency inputs.

References

Achuthan S, Canavier CC. Phase-resetting curves determine synchronization, phase locking, and clustering in networks of neural oscillators. Journal of Neuroscience. 2009; 29(16):5218–5233.

Acker CD, Kopell N, White JA. Synchronization of strongly coupled excitatory neurons: relating network behavior to biophysics. Journal of computational neuroscience. 2003; 15(1):71–90.

Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. Proceedings of the National Academy of Sciences. 2001; 98(23):13367–13372.

Akam TE, Kullmann DM. Efficient "communication through coherence" requires oscillations structured to minimize interference between signals. PLoS computational biology. 2012; 8(11):e1002760.

Arnal LH, Giraud AL. Cortical oscillations and sensory predictions. Trends in cognitive sciences. 2012; 16(7):390–398.

Atallah BV, Scanziani M. Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. Neuron. 2009; 62(4):566–577.

Aydore S, Pantazis D, Leahy RM. A note on the phase locking value and its properties. NeuroImage. 2013; 74:231–244. doi: 10.1016/j.neuroimage.2013.02.008.

Baese-Berk MM, Heffner CC, Dilley LC, Pitt MA, Morrill TH, McAuley JD. Long-term temporal tracking of speech rate affects spoken-word recognition. Psychological Science. 2014; 25(8):1546–1553.

Bosker HR, Ghitza O. Entrained theta oscillations guide perception of subsequent speech: behavioural evidence from rate normalisation. Language, Cognition and Neuroscience. 2018; 33(8):955–967.

Brown M, Salverda AP, Dilley LC, Tanenhaus MK. Expectations from preceding prosody influence segmentation in online sentence processing. Psychonomic bulletin & review. 2011; 18(6):1189–1196.

Brown M, Salverda AP, Dilley LC, Tanenhaus MK. Metrical expectations from preceding prosody influence perception of lexical stress. Journal of Experimental Psychology: Human Perception and Performance. 2015; 41(2):306.

Buzsaki G. Rhythms of the Brain. Oxford University Press; 2006.

Canavier CC, Achuthan S. Pulse coupled oscillators and the phase resetting curve. Mathematical biosciences. 2010; 226(2):77–96.

Cannon J, Kopell N. The leaky oscillator: Properties of inhibition-based rhythms revealed through the singular phase response curve. SIAM Journal on Applied Dynamical Systems. 2015; 14(4):1930–1977.

Carracedo LM, Kjeldsen H, Cunnington L, Jenkins A, Schofield I, Cunningham MO, Davies CH, Traub RD, Whittington MA. A neocortical delta rhythm facilitates reciprocal interlaminar interactions via nested theta rhythms. Journal of Neuroscience. 2013; 33(26):10750–10761.

Chandrasekaran C, Trubanova A, Stillittano S, Caplier A, Ghazanfar AA. The natural statistics of audiovisual speech. PLoS computational biology. 2009; 5(7):e1000436.

Christiansen MH, Chater N. The Now-or-Never bottleneck: A fundamental constraint on language. Behavioral and Brain Sciences. 2016; 39.

Dilley LC, Mattys SL, Vinke L. Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. Journal of Memory and Language. 2010; 63(3):274–294.

Dilley LC, Pitt MA. Altering context speech rate can cause words to appear or disappear. Psychological Science. 2010; 21(11):1664–1670.

Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. Nature neuroscience. 2016; 19(1):158.

Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D. Temporal modulations in speech and music. Neuroscience & Biobehavioral Reviews. 2017; .

Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. Neuroimage. 2014; 85:761–768.

Drullman R, Festen JM, Plomp R. Effect of reducing slow temporal modulations on speech reception. The Journal of the Acoustical Society of America. 1994; 95(5):2670–2680.

Elliott TM, Theunissen FE. The modulation transfer function for speech intelligibility. PLoS computational biology. 2009; 5(3):e1000302.

Ermentrout B. Type I membranes, phase resetting curves, and synchrony. Neural computation. 1996; 8(5):979–1001.

Ermentrout B, Pascal M, Gutkin B. The effects of spike frequency adaptation and negative feedback on the synchronization of neural oscillators. Neural computation. 2001; 13(6):1285–1310.

Ermentrout GB. n: m Phase-locking of weakly coupled oscillators. Journal of Mathematical Biology. 1981; 12(3):327–342.

Ghitza O. Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. Frontiers in psychology. 2011; 2:130.

Ghitza O. On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. Frontiers in psychology. 2012; 3:238.

Ghitza O. The theta-syllable: a unit of speech information defined by cortical function. Frontiers in psychology. 2013; 4:138.

Ghitza O. Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech. Frontiers in psychology. 2014; 5:652.

Ghitza O. Neuronal oscillations in decoding time-compressed speech. The Journal of the Acoustical Society of America. 2016; 139(4):2190–2190.

Ghitza O. Acoustic-driven delta rhythms as prosodic markers. Language, Cognition and Neuroscience. 2017; 32(5):545–561.

Ghitza O, Greenberg S. On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. Phonetica. 2009; 66(1-2):113–126.

Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nature neuroscience. 2012; 15(4):511.

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS biology. 2013; 11(12):e1001752.

Gutfreund Y, Segev I, et al. Subthreshold oscillations and resonant frequency in guinea-pig cortical neurons: physiology and modelling. The Journal of physiology. 1995; 483(3):621–640.

Henry MJ, Herrmann B, Obleser J. Entrained neural oscillations in multiple frequency bands comodulate behavior. Proceedings of the National Academy of Sciences. 2014; 111(41):14935–14940.

Hertrich I, Dietrich S, Trouvain J, Moos A, Ackermann H. Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal. Psychophysiology. 2012; 49(3):322–334.

Hirst D, Di Cristo A. Intonation systems: a survey of twenty languages. Cambridge University Press; 1998.

Hu H, Vervaeke K, Storm JF. Two forms of electrical resonance at theta frequencies, generated by M-current, h-current and persistent Na+ current in rat hippocampal pyramidal cells. The Journal of physiology. 2002; 545(3):783–805.

Huggins AWF. Distortion of the temporal pattern of speech: Interruption and alternation. The Journal of the Acoustical Society of America. 1964; 36(6):1055–1064.

Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud AL. Speech encoding by coupled cortical theta and gamma oscillations. Elife. 2015; 4.

Kayser C, Wilson C, Safaai H, Sakata S, Panzeri S. Rhythmic auditory cortex activity at multiple timescales shapes stimulus–response gain and background firing. Journal of Neuroscience. 2015; 35(20):7750–7762.

Kleinfeld D, Ahissar E, Diamond ME. Active sensation: insights from the rodent vibrissa sensorimotor system. Current opinion in neurobiology. 2006; 16(4):435–444.

Kleinfeld D, Deschenes M, Ulanovsky N. Whisking, sniffing, and the hippocampal θ-rhythm: a tale of two oscillators. PLoS biology. 2016; 14(2):e1002385.

Kopell N, Ermentrout G. Mechanisms of phase-locking and frequency control in pairs of coupled neural oscillators. Handbook of dynamical systems. 2002; 2:3–54.

Lakatos P, Musacchia G, O'Connel MN, Falchier AY, Javitt DC, Schroeder CE. The spectrotemporal filter mechanism of auditory selective attention. Neuron. 2013; 77(4):750–761.

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. Journal of neurophysiology. 2005; 94(3):1904–1911.

Lewis AG, Bastiaansen M. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. Cortex. 2015; 68:155–168.

Luce PA, CONOR T. 24 Spoken Word Recognition: The Challenge of Variation. The handbook of speech perception. 2005; p. 591.

Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron. 2007; 54(6):1001–1010.

Mai G, Minett JW, Wang WSY. Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. Neuroimage. 2016; 133:516–528.

Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition. 1987; 25(1-2):71–102.

McFarland WL, Teitelbaum H, Hedges EK. Relationship between hippocampal theta activity and running speed in the rat. Journal of comparative and physiological psychology. 1975; 88(1):324.

Miller GA, Licklider JC. The intelligibility of interrupted speech. The Journal of the Acoustical Society of America. 1950; 22(2):167–173.

Morillon B, Schroeder CE. Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. Annals of the New York Academy of Sciences. 2015; 1337(1):26–31.

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

Ohala JJ. The temporal regulation of speech. Auditory analysis and perception of speech. 1975; p. 431–453.

Peelle JE, Gross J, Davis MH. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. Cerebral cortex. 2012; 23(6):1378–1387.

Pefkou M, Arnal LH, Fontolan L, Giraud AL. θ-Band and β-Band Neural Activity Reflects Independent Syllable Tracking and Comprehension of Time-Compressed Speech. Journal of Neuroscience. 2017; 37(33):7930–7938.

Penn LR, Ayasse ND, Wingfield A, Ghitza O. The possible role of brain rhythms in perceiving fast speech: Evidence from adult aging. The Journal of the Acoustical Society of America. 2018; 144(4):2088–2094.

Räsänen O, Doyle G, Frank MC. Pre-linguistic segmentation of speech into syllable-like units. Cognition. 2018; 171:130–150.

Riecke L, Formisano E, Herrmann CS, Sack AT. 4-Hz transcranial alternating current stimulation phase modulates hearing. Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation. 2015; 8(4):777–783.

Riecke L, Formisano E, Sorger B, Başkent D, Gaudrain E. Neural Entrainment to Speech Modulates Speech Intelligibility. Current Biology. 2017; .

Riecke L, Sack AT, Schroeder CE. Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. Current Biology. 2015; 25(24):3196–3201.

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. Phil Trans R Soc Lond B. 1992; 336(1278):367–373.

Rotstein HG. Spiking resonances in models with the same slow resonant and fast amplifying currents but different subthreshold dynamic properties. Journal of computational neuroscience. 2017; 43(3):243–271.

Rotstein HG, Nadim F. Frequency preference in two-dimensional neural models: a linear analysis of the interaction between resonant and amplifying currents. Journal of computational neuroscience. 2014; 37(1):9–28.

Rotstein HG, Pervouchine DD, Acker CD, Gillies MJ, White JA, Buhl EH, Whittington MA, Kopell N. Slow and fast inhibition and an H-current interact to create a theta rhythm in a model of CA1 interneuron network. Journal of neurophysiology. 2005; 94(2):1509–1518.

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

Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A. Neuronal oscillations and visual amplification of speech. Trends in cognitive sciences. 2008; 12(3):106–113.

Selkirk EO. The role of prosodic categories in English word stress. Linguistic inquiry. 1980; 11(3):563–605.

Sherfey JS, Ardid S, Hass J, Hasselmo ME, Kopell NJ. Flexible resonance in prefrontal networks with strong feedback inhibition. PLoS computational biology. 2018; 14(8):e1006357.

Sherfey JS, Soplata AE, Ardid S, Roberts EA, Stanley DA, Pittman-Polletta BR, Kopell NJ. DynaSim: a MATLAB Toolbox for neural modeling and simulation. Frontiers in neuroinformatics. 2018; 12:10.

Shin D, Cho KH. Recurrent connections form a phase-locking neuronal tuner for frequency-dependent selective communication. Scientific reports. 2013; 3:2519.

Stevens KN. Features in speech perception and lexical access. The handbook of speech perception. 2005; p. 125–155.

Stilp CE, Kiefte M, Alexander JM, Kluender KR. Cochlea-scaled spectral entropy predicts rate-invariant intelligibility of temporally distorted sentences. The Journal of the Acoustical Society of America. 2010; 128(4):2112–2126.

Ten Oever S, Sack AT. Oscillatory phase shapes syllable perception. Proceedings of the National Academy of Sciences. 2015; 112(52):15833–15837.

Teng X, Tian X, Doelling K, Poeppel D. Theta band oscillations reflect more than entrainment: behavioral and neural evidence demonstrates an active chunking process. European Journal of Neuroscience. 2017; .

Traub RD, Wong RK, Miles R, Michelson H. A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. Journal of Neurophysiology. 1991; 66(2):635–650.

Tsai TYC, Choi YS, Ma W, Pomerening JR, Tang C, Ferrell JE. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. Science. 2008; 321(5885):126–129.

Wilsch A, Neuling T, Herrmann CS. Envelope-tACS modulates intelligibility of speech in noise. bioRxiv. 2017; p. 097576.

Wilsch A, Neuling T, Obleser J, Herrmann CS. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. NeuroImage. 2018; 172:766–774.

Yang Lc. Duration and Pauses as Boundary-Markers in Speech: A Cross-Linguistic Study. In: *Eighth Annual Conference of the International Speech Communication Association*; 2007. .

Yang X, Shen X, Li W, Yang Y. How listeners weight acoustic cues to intonational phrase boundaries. PloS one. 2014; 9(7):e102166.

Zhou Y, Vo T, Rotstein HG, McCarthy MM, Kopell N. M-Current Expands the Range of Gamma Frequency Inputs to Which a Neuronal Target Entrains. The Journal of Mathematical Neuroscience. 2018; 8(1):13.

Zoefel B, Archer-Boyd A, Davis MH. Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. Current Biology. 2018; .

Zoefel B, VanRullen R. The role of high-level processes for oscillatory phase entrainment to speech sound. Frontiers in human neuroscience. 2015; 9:651.