Dynamics of the auditory continuity illusion

Qianyi Cao[†], Noah Parks[†] and Joshua H. Goldwyn^{*}

Swarthmore College, Department of Mathematics and Statistics, Swarthmore, PA, USA

Correspondence*:
Joshua H. Goldwyn
jhgoldwyn@gmail.com

[†] These authors contributed equally to this work.

2 ABSTRACT

3 Illusions give intriguing insights into perceptual and neural dynamics. In the auditory continuity
4 illusion, two brief tones separated by a silent gap may be heard as one continuous tone if a
5 noise burst with appropriate characteristics fills the gap. This illusion probes the conditions
6 under which listeners link related sounds across time and maintain perceptual continuity in the
7 face of sudden changes in sound mixtures. Conceptual explanations of this illusion have been
8 proposed, but its neural basis is still being investigated. In this work we provide a dynamical
9 systems framework, grounded in principles of neural dynamics, to explain the continuity illusion.
10 We construct an idealized firing rate model of a neural population and analyze the conditions
11 under which firing rate responses persist during the interruption between the two tones. First, we
12 show that sustained inputs and hysteresis dynamics (a mismatch between tone levels needed
13 to activate and inactivate the population) can produce continuous responses. Second, we show
14 that transient inputs and bistable dynamics (coexistence of two stable firing rate levels) can also
15 produce continuous responses. Finally, we combine these input types together to obtain neural
16 dynamics consistent with two requirements for the continuity illusion as articulated in a well-known
17 theory of auditory scene analysis: sustained responses occur if noise provides sufficient evidence
18 that the tone continues and if there is no evidence of discontinuities between the tones and noise.
19 By grounding these notions in a quantitative model that incorporates elements of neural circuits
20 (recurrent excitation, and mutual inhibition, specifically), we identify plausible mechanisms for the
21 continuity illusion. Our findings can help guide future studies of neural correlate of this illusion
22 and inform development of more biophysically-based models of the auditory continuity illusion.

23 **Keywords:** auditory scene analysis, bistability, continuity illusion, computational neuroscience, hysteresis, neural dynamics

1 INTRODUCTION

24 How do listeners in crowded and noisy environments create stable auditory streams in the face of
25 interruptions and “background” noise? How do listeners identify the stops and starts of overlapping
26 and interwoven sounds to correctly parse an auditory scene? Answering these questions is fundamental
27 to understanding auditory perception and neural processing of sounds. A perceptual illusion that sheds
28 light on dynamic processing of multiple sounds is the auditory continuity illusion [1] (also called temporal
29 induction [2]). The continuity illusion can be elicited when noise interrupts a variety of sounds including
30 tones, frequency glides, sentences [1, 2], and sound textures [3]. The common aspect of this illusion is that,
31 when the noise is sufficiently loud and shares spectral content with the interrupted signal, listeners perceive
32 a continuous, uninterrupted sound. This illusion reveals a tendency for the auditory system to maintain
33 perceptual continuity when confronted with sudden changes in the auditory scene and to sustain perception

of sounds that are present prior to some masking or distracting sound. The continuity illusion has been thoroughly studied since its discovery [4, 5].

Figure 1. Illustration of the continuity illusion. (A) A tone interrupted by a silent gap (top panel) is perceived (correctly) as two tones discontinuous in time (bottom panel). (B) If a weak noise burst is inserted into the gap between tones, the tone remains perceived as discontinuous. (C) A sufficiently loud noise that shares frequency content with the tone can induce the illusion that the tone is uninterrupted and persists through the noise-filled interruption.

The version of the continuity illusion we investigate is an interrupted tone that can be perceived as continuous if the interruption (a short interval in which no tone is presented) is filled with broadband noise, as depicted in Fig. 1. Results of listening experiments inform conceptual models for the illusion such as Bregman’s theories of auditory scene analysis [1]. Fundamental questions remain about the perceptual origin and neural basis for the illusion (viz. whether it is a cortical phenomenon [6, 7, 8] or created at subcortical stations [9]) and whether it is required that peripheral responses signal discontinuities in the sound [10]). Perhaps due to these uncertainties, there have been few efforts to model how the dynamic activity of neural populations can generate the continuity illusion. Previous works point toward several possible neural mechanisms that can give rise to the illusion: feedforward intra-cortical connections [6], nonlinear dynamic self-excitation [11], and short-term synaptic plasticity [12]). These few and disparate studies are motivation for further model-based investigations of the continuity illusion.

Our goal is to identify dynamical mechanisms that implement fundamental principles of the illusion. We are informed by the work of Petkov and colleagues who obtained evidence that two types of neural responses participate in the continuity illusion: sustained responses that signal ongoing sounds and transient responses that signal acoustic edges (onsets and offsets) [8]. While they specifically identified these response types in auditory cortex, we view these as common response motifs that are found throughout the auditory pathway. (see [13] for a review of offset auditory neurons, e.g.). In addition, we seek to connect Bregman’s principles, gained over decades of careful observation, to fundamental features of neural dynamics. Specifically, we present dynamical explanations for two of Bregman’s “rules” that define the circumstances under which the illusion will occur [1]. We paraphrase these rules as:

Sufficiency of Evidence Rule: There must be some neural activity during the interruption that is indistinguishable from what would have occurred if the tone had continued through the noise-filled interruption.

No Discontinuity Rule: There must be no evidence that the tone shuts off during the noise-filled interruption.

We use a nonlinear, dynamical firing rate model [14] with sustained inputs to implement the Sufficiency of Evidence rule. The dynamical principle at work is hysteresis: the interrupting noise (a broadband sound) provides a partial amount of excitatory input to an idealized neural population. Recurrent excitation in the

population enables the noise to maintain firing activity in an already activated population, even though the noise alone cannot activate an inactive population. We then adjust excitability in the model so that it has two stable states that coexist in the absence of sustained inputs (bistability). We use this configuration to show that transient inputs at tone onsets and offsets implement the No Discontinuity rule. Continuity occurs when noise suppresses the offset response at the end of the first tone. Without this sufficiently strong offset response, the population remains active despite the absence of any ongoing input during the interruption between tones. Finally, we configure the model to receive both input types. In this setting, the model utilizes both hysteresis and bistability to create dynamics consistent with the continuity illusion. This formulation of using persistent neural activity as an indicator of perceptual state, instantiated as an attractor in a dynamical system, resembles ideas common to studies of the neural basis for working memory [15].

We illustrate our dynamical explanation for the continuity illusion with simulation result and we emphasize, throughout, analytical and geometric insights gained by computing equilibrium solutions to the nonlinear differential equation that governs the firing rate dynamics. By identifying principles of neural dynamics that are consistent with the continuity illusion, our work can guide future modeling studies that seek to simulate the illusion with more biophysically-detailed neural networks and inform studies that seek to identify the neural circuits and mechanisms that create the illusion.

2 MATERIAL AND METHODS

2.1 Firing rate model of a neural population

We model sound-evoked activity of a neural population using a firing rate equation with recurrent excitation [14]. The differential equation that describes the dynamics of the neural activity is

$$\tau x' = -x + f(a_E x + I(t)) \quad (1)$$

where $x(t)$ is the firing rate variable that takes dimensionless values between 0 (population inactive) and 1 (population active). The value of x can be interpreted in a mean-field sense as the proportion of active neurons in a population or as the instantaneous probability of firing for neurons in the population. We suppose that an activity level near one for this population signals perception of a tone at some frequency. The parameter τ is a time constant of the firing rate dynamics and a_E is the strength of recurrent excitation within the population. The transfer function f is the sigmoidal nonlinearity

$$f(u) = \frac{1}{1 + e^{-(u-m)/k}}. \quad (2)$$

Since we are working with dimensionless quantities, we use $k = 1$ throughout without loss of generality. We explain our choices for the half-maximum parameter m in more detail below.

The input term $I(t)$ is composed of for components:

$$I(t) = I_{sustain} + I_{onset} - I_{offset} - I_{inhib}. \quad (3)$$

The first term is a sustained tone-driven excitatory responses ($I_{sustain}$, constant for the duration of the tone). These second and third terms are transient, exponentially-decaying events triggered by tone onsets and offsets. The minus sign for I_{offset} indicates that offset responses are inhibitory inputs to the population. In

some cases, there is also a noise-driven inhibitory term I_{inhib} that is a sustained response (constant for the duration of the noise). The classification of inputs as either sustained or transient, and the description of their dynamics as either constant or exponentially-decaying is a caricature of actual inputs motivated by observations by Petkov et al. [8]. Those authors proposed that sustained and transient response types in primary auditory cortex are possible neural correlates of the continuity illusion.

For responses to tones alone (no noise), the strengths of the tone-driven inputs are proportional to a dimensionless tone level parameter I_T . By convention, we parameterize all models so that $I_T = 1$ is the threshold for tone perception (in the absence of noise). We also restrict I_T so it does not exceed a maximum tone level that we set (arbitrarily) at $I_{T,max} = 5$. The effect of noise varies by input type. Sustained inputs can be thought to represent the ongoing energy in the frequency-channel to which the population is tuned. Since noise is a broadband signal, we allow noise inputs to provide partial input to $I_{sustain}$. Transient inputs can be thought to represent the salience of the onsets and offsets of a tone. We suppose noise degrades the “sharpness” of these acoustic edges and thus noise reduces I_{onset} and I_{offset} if noise is present at the instants of tone onset or offset. Finally, in some cases, we suppose there is a noise-driven pool that directly inhibits the tone population and that, in turn, receives inhibition from the tone population. For analytical convenience, we suppose this noise-driven inhibitory pool operates on a time-scale faster than the dynamics of x , and thus we let I_{inhib} evolve instantaneously to a value determined by noise strength. We use I_N to represent a dimensionless measure of noise level that takes value between 0 (threshold for noise perception) and $I_{N,max} = 10$ (arbitrarily chosen maximum noise level). From these considerations we have

$$\begin{aligned} I_{sustain} &= I_T + \alpha I_N \\ I_{onset} &= \gamma_{on} [I_T - \beta I_N]_+ e^{-(t-t_{on})/\tau} H(t - t_{on}) \\ I_{offset} &= \gamma_{off} [I_T - \beta I_N]_+ e^{-(t-t_{off})/\tau} H(t - t_{off}) \\ I_{inhib} &= a_I I_N (1 - x) \end{aligned} \tag{4}$$

The parameters α and β are between 0 and 1 and represent the effects of noise on the tone-driven responses, as described above. The parameter γ_{on} scales the onset response so that tone threshold is always $I_T = 1$. We detail the calculation of γ_{on} below. The parameter γ_{off} similarly scales the offset response. We choose the decay time constant τ for transient responses to match the time scale of the firing rate dynamics. The notation $[\cdot]_+$ denotes rectification (values bound below at a floor of 0). The function H is the Heaviside function (also known as a step function) that is 0 if its argument is less than 0 and 1 if its argument is greater than 0. We use it here to indicate that the transient onset and offset responses begin at tone onset (t_{on}) and tone offset (t_{off}), respectively. The parameter a_I is the strength of the noise-driven inhibition. The multiplicative factor $(1 - x)$ in I_{inhib} implements mutual inhibition between the tone population and the noise-driven pool whose activity is represented by I_{inhib} since increases in the firing rate x reduce the strength of this noise-driven inhibitory input.

2.2 Stimulus configurations

We perform simulations and analyzing activation dynamics for three stimulus configurations: tone only, tone masked by noise, and tones interrupted by a noise-filled gap. In the masking configuration, tone and noise are presented simultaneously so that the noise impacts sustained and transient responses throughout the duration of the tone. In the interrupted tones configuration, the tone is shut off during the interruption

129 between ($I_T = 0$) and noise impacts the offset response of the first tone and the onset response of the
130 second tone.

131 As mentioned above, all models are parameterized so that $I_T = 1$ is the threshold for firing rate activation
132 in the tone only case. In the masking case, we analyze activation dynamics and determine the smallest
133 noise level at which a tone at a given level can overcome the effects of noise and activate the population.
134 This defines the masking threshold $M(I_T)$. Similarly, for the interrupted tones, we determine the noise
135 level (for a given tone level) at which the population remains active during the gap between tones. This
136 defines the continuity threshold $C(I_T)$.

137 In simulations we use 1-second long inputs for tones (in tone only case) and tone and noise (in masking
138 case). To simulate the continuity illusion, we present two tones at the same level and fill the interrupting
139 gap with noise. The total duration of the continuity input is 2.5 seconds (tones are each 1 second long,
140 the noise-filled gap is 0.5 seconds long). To be clear, tones and noise are represented by I_T and I_N in our
141 model, we do not include sinusoidal waveforms for tones or stochastic processes for noise.

142 2.3 Numerical simulations

143 All calculations were carried out using the scientific computing software MATLAB (The MathWorks,
144 Inc.). The firing rate differential equation (Eq. 1) was solved numerically using `ode15s`. Simulation code is
145 available for download and use at <https://github.com/jhgoldwyn/ContinuityIllusion>.
146 Nonlinear equations (to compute equilibrium states, for example) were solved using root-finding functions
147 in Matlab such as `fzero` and `fsolve`.

3 RESULTS

148 3.1 Model classification by firing rate equilibria

149 We begin with a general analysis of the equilibrium states of the model for the case of sustained tone
150 input only ($I_{onset} = I_{offset} = 0$ and $I_N = 0$). These calculations inform our parameter choices and are a
151 starting point for subsequent analyses. Setting $x' = 0$ and $I(t) = I_T$ in Eq. 1) and solving for I_T we find
152 the equilibrium relation

$$I_T(x) = m - \ln\left(\frac{1}{x} - 1\right) - a_E x. \quad (5)$$

153 The parameters a_E and m determine the shape of $I_T(x)$. We show three representative examples in
154 Fig. 2A, with equilibrium firing rates plotted on the vertical axis and I_T plotted on the horizontal axis. The
155 key features of these curves are: whether they are S-shaped with left and right “knees” and, if so, the values
156 of I_T at these knees. Recall that the firing rate variable x takes values between zero and one. We interpret
157 equilibrium solutions $x \approx 0$ as inactive states (no perception of tone) and equilibrium solutions $x \approx 1$ as
158 active states (perception of tone).

159 We summarize the possible scenarios in Table 1 and show how they partition the a_E - m plane in Fig. 2B.
160 The scenarios that interest us must satisfy three criteria. First, to clearly distinguish between active and
161 inactive states, the x - I_T curve must be S-shaped. This rules out Region I. Second, the left knee must be
162 located at an I_T value less than the maximum tone strength ($I_{T,max} = 5$) so that activation is possible.
163 This rules out Region II. Lastly, the right knee must be located at a positive I_T value so that inactivation is

Region	Left knee	Right knee
I	none	none
II	$> I_{T,max}$	$> I_{T,max}$
IIIb	$(0, I_{T,max})$	$> I_{T,max}$
IIIa	$(0, I_{T,max})$	$(0, I_{T,max})$
IV	< 0	> 0
V	< 0	< 0

Table 1. Regions of a_E - m parameter space classified by positions of left and right knees of x - I_T equilibrium firing rate curves. Regions of interest for this study are II and III. Other regions are ruled out because inactive and active states are indistinguishable (I), active state is inaccessible (II), or inactive state is inaccessible (V).

possible. This rules out Region V, in which the population could remain active at all times, even without inputs.

Figure 2. Partition of parameter space by shape of firing rate equilibrium curves. (A) Equilibrium firing rates for tone-only inputs ($I_N = 0$). Two stable branches (solid) are separated by a branch of unstable equilibria (dotted). Gray line at right indicates the maximum tone level used in simulations and analysis ($I_{T,max} = 5$). (B) Parameter space for parameters in Eqs. 1 and 2. Colored dots correspond to the curves in (A). Parameter space is partitioned according position of left and right knees of the firing rate equilibrium curves (see Table 1 for details).

As detailed in Table 1, the number and locations of the knees of the equilibrium curve delineate these regions. The knees are saddle node bifurcation points in the firing rate dynamics (points at which stable and unstable equilibria appear or disappear). They are located at critical points of $I_T(x)$, so we identify these points by taking a derivative of Eq. 5 with respect to x and setting the resulting expression to 0. The x values at the left and right knees are

$$\begin{aligned} x_L &= \frac{1}{2} \left(1 + \sqrt{1 - 4/a_E} \right) \\ x_R &= \frac{1}{2} \left(1 - \sqrt{1 - 4/a_E} \right). \end{aligned} \quad (6)$$

We compute the value of I_T at the knees by plugging these values into Eq. 5. In particular, we parameterize a_E and m by specifying $I_T(x_L)$ and $I_T(x_R)$ and solving the resulting two nonlinear equations given by Eq. 5. No real solutions exist for $a_E < 4$, thus $a_E = 4$ marks the boundary of models without S-shaped equilibrium curves (region I in Fig. 2B). The regions of interest for this study (III and IV) by the positions of the knees in the equilibrium curve. In particular, region IV consists of (a_E, m) parameter pairs for which the model is bistable in the absence of any inputs (the left knee is at a negative I_T value and the right knee is at a positive I_T value). Region III consists of models that are monostable with no inputs (left knee is at a positive I_T value). We distinguish between two subregions of Region III. In Region IIIa, I_T at the right knee is less than the maximum tone value, and thus sustained inputs alone can activate the neural population. In contrast, in Region IIIb the I_T value at the right knee is larger than the maximum tone strength and thus models in this region can only be activated by a combination of sustained and transient inputs.

From these considerations, we choose three models (one from each region of interest) and use these in all further analysis and simulations.

Model 1 ($a_E = 5.9, m = 3.6$, in Region IIIa). The left knee is at $I_T = 0.2$ and the right knee is at $I_T = 1$. As we describe below, the essential dynamical feature of this model is hysteresis: in response to sustained inputs the system requires a higher tone level to activate than deactivate (the S-shape of the equilibrium curve creates a mismatch between the two knees). We use sustained inputs only for this model ($I_{onset} = I_{offset} = 0$ in Eq. 3).

Model 2 ($a_E = 10.5, m = 5.2$, in Region IV). The left knee is at $I_T = -2$ and the right knee is at $I_T = 2$. As we describe below, the essential dynamical feature of this model is bistability. We use transient inputs to move the system between active and inactive states and do not include sustained inputs ($I_{sustain} = I_{inhib} = 0$ in Eq. 3).

Model 3 ($a_E = 12.7, m = 9.5$, in Region IIIb). The left knee is at $I_{sust} = 0.2$ and the right knee is at $I_{sust} = 6$. Sustained inputs alone cannot activate models in Region IIIb, so we use a combination of transient and sustained inputs for this model (all input terms in Eq 3 are non-zero).

Next, we show how each of these three models can demonstrate firing activity consistent with masking and the continuity illusion. We also compute masking and continuity thresholds analytically and explore effects of key parameters. We begin with Model 1 and show how it utilizes sustained inputs and the hysteresis effect to create firing rate dynamics consistent with the continuity illusion. We then show that Model 2, which uses only transient inputs but is set at a bistable point, can also produce the continuity illusion. Since sustained inputs signal (ongoing) evidence of a tone and transient inputs signal tone onsets and offsets, we view Model 1 as an implementation of Bregman's No Discontinuity rule and Model 2 as an implementation of the No Discontinuity rule. Lastly, we show Model 3, which uses a combination of input types (sustained and transient), implements both rules for the illusion. Firing activity in Model 3 can persist through an interruption between tones because noise has two effects: to contribute excitatory inputs during the interruption (sufficient evidence), and to prevent the offset response from inactivating the population (no discontinuity).

3.2 Model 1: Sustained inputs and hysteresis dynamics implement the Sufficiency of Evidence Rule

3.2.1 Response dynamics to tone-only inputs

Recall that Model 1 includes sustained inputs only ($I_{onset} = I_{offset} = 0$). The equilibrium solutions for this model for tone inputs only ($I_N = 0$) are shown in Fig. 3A1 with I_T as the bifurcation parameter

(horizontal axis). When there is no tone input ($I_T = 0$), this system has a single stable equilibrium in the inactive state. For stronger tone inputs, the system passes through a saddle point bifurcation point at which a second stable equilibrium is created in the active state. Activation of the population from rest requires the tone to be larger, namely that $I_T > I_R$ where I_R is the tone level at the right knee of the equilibrium curve. At this second saddle node bifurcation point, the inactive state is abolished and only the active state remains as the unique stable and globally attracting fixed point for the system. Responses to subthreshold and suprathreshold tone inputs are shown in Fig. 3A2.

The feature of this model that is essential in our study of the continuity illusion (discussed below) is that it exhibits hysteresis dynamics. By this we mean that the tone level that activates an inactive population is larger than the tone level that maintains an already active population in the active state. Hysteresis is seen geometrically in the S-shaped x - I_T equilibrium curve (Fig. 3A1). The activation threshold is the tone strength at the right knee. The deactivation threshold – the minimum tone level that maintains activity – is the tone strength at the left knee. For Model 1, we positioned these knees at $I_T = 1$ (activation) and $I_T = 0.2$ (deactivation). As an example of the hysteresis effect: tone input with $I_T = 0.5$ does not activate x from rest (blue curve in Fig. 3A2) but it would maintain x at a level near 1 if x were active prior to this input (not shown, but notice the upper branch of equilibria in Fig. 3A1 extends to I_T values less than 0.5).

3.2.2 Response dynamics to tone masked by noise

The two effects of noise are that it provides a “partial” input that enters as the additive term αI_N in $I_{sustain}$, and that it produces noise-driven inhibition through the term $I_{inhib} = a_I I_N(1 - x)$ (recall Eq. 4). The effect of the inhibitory term is to shift the right knee of the equilibrium curve to larger I_T levels (Fig. 3B1, compare black curve with noise to gray curve without noise). That is, the threshold for activation increases with I_N , as desired for noise to have a masking effect. As a demonstration, a tone level of $I_T = 1.5$ would activate the population from rest in the absence of noise, but with $I_N = 1$ this input does not activate the population (red curve in Fig. 4B2).

The equilibrium solutions for this model, using I_T as the bifurcation parameter and now including the effect of noise, are

$$I_T(x, I_N) = m - \ln\left(\frac{1}{x} - 1\right) - a_E x + a_I I_N(1 - x) - \alpha I_N. \quad (7)$$

The critical points at which stable equilibria are created and abolished are the knees of the S-shaped curve. We obtain results similar to Eq. 6, but now with noise included:

$$\begin{aligned} x_L(I_N) &= \frac{1}{2} \left(1 + \sqrt{1 - 4/(a_E + a_I I_N)} \right) \\ x_R(I_N) &= \frac{1}{2} \left(1 - \sqrt{1 - 4/(a_E + a_I I_N)} \right). \end{aligned} \quad (8)$$

The $x_R(I_N)$ point locates the threshold at which a tone activates a population from rest, in the presence of noise. The threshold for masking, then, is the noise level that solves $I_T = I_R(I_N)$. Any value of I_T smaller than this critical value would fail to activate the population due to the noise-driven inhibition. We denote this masking threshold tone level as $M(I_N)$. It depends nonlinearly on I_N because of the dependence of x_R on I_N . For Model 1, however, x_R is relatively constant with respect to I_N and we find masking threshold

Figure 3. Dynamics of Model 1 responses to tone and noise inputs. (A) Tone inputs (no noise) with tone level $I_T = 0.5$ (blue) and $I_T = 1.5$ (red). Equilibrium firing rates shown in panel A1 and time-courses of firing rate variable $x(t)$ in panel A2. Firing rate equilibrium in A1 (tone only) is re-plotted in gray in B1 and C1. (B) Simultaneous tone and noise inputs (masking condition) with $I_N = 1$. Noise shifts right knee of equilibrium curve to larger I_T levels (B1) and prevents activation by tone inputs ($I_T = 0.5$ and $I_T = 1.5$ shown). (C) Interrupted tone with noise-filled gap (continuity condition) with $I_T = 1.5$. Noise shifts left knee of equilibrium curve to smaller I_T levels (C1). Continuity occurs if left knee crosses $I_T = 0$ axis. Time-courses of firing rate variable in C2 for no noise (blue) and noise above continuity threshold ($I_N = 8$, red). (D) Masking threshold and continuity thresholds. (E) Masking and continuity thresholds at maximum tone level, for varying parameter a_I and α parameter values. $C(I_{T,max})$ shown as a color map, contour lines show $M(I_{T,max})$ (values labeled above contours). Parameter values used in panels (A - D): $a_I = 1.124$, $\alpha = 0.168$, shown as black dot in E.

as a function of I_T is approximately linear (gray curve in Fig. 3D). The slope of this nearly linear relation can be approximated by $[a_I(1 - x_R(0)) - \alpha]^{-1}$ (found by taking the derivative of I_T with respect to I_N , and neglecting any change in x_R with respect to I_N). This approximation shows the opposing effects of noise on masking threshold: $M(I_N)$ increases with increasing α (the amount of excitatory noise input) and decreases with increasing a_I (the amount of inhibitory noise-driven input), see Fig. 3H. This relation also imposes a constraint on model parameters. We must have $\alpha < a_I(1 - x_R(0))$. If this condition is not satisfied, the excitatory effect of noise (the αI_N term in $I_{sustain}$) would dominate the inhibitory effect of noise (I_{inhib}) and masking would not be possible.

3.2.3 Response dynamics to tones interrupted by noise

The x_L point locates the threshold for inactivation in the presence of noise. A population in the active state will remain active during the gap between tones if the noise strength causes $I_T(x_L)$ to cross over to negative values, see Fig. 3C1. Thus, we calculate the continuity threshold equation by solving (numerically) the root-finding problem $I_T(x_L, I_N) = 0$, where x_L is also a function of noise level (Eq. 8). Observe that, since $I_T = 0$ during the gap between tones, the continuity threshold is constant with respect to tone level, as shown by the horizontal black line in Fig. 3D.

The masking and continuity thresholds for this model are separate. This imposes additional constraints on our parameter choices. Specifically, in accordance with the hypothesis that the continuity illusion is a compensation for masking [2, 16], we require that continuity can only occur at noise levels at least as high as the masking threshold. Additionally, we are only interested in parameter sets for which masking and continuity can both be achieved (continuity and masking threshold must not exceed $I_{N,max} = 10$, even for the strongest tones ($I_{T,max} = 5$). A view of the $a_I - \alpha$ parameter region that satisfies these requirements is shown in Fig. 3H, with labeled contour lines indicating the corresponding masking and continuity thresholds at the maximum tone level.

3.3 Model 2: Transient inputs and bistable dynamics implement the No Discontinuity Rule

3.3.1 Response dynamics to tone-only inputs

Model 2 receives transient inputs only ($I_{sustain} = I_{inhib} = 0$). These transient inputs occur at tone onsets and offsets and signal discontinuities at the “edges” of a tone. To analyze dynamics of this model at tone onset we find it useful to formulate it as a system of two ordinary differential equations:

$$\begin{aligned}\tau x' &= -x + f(a_E x + \gamma_{on} s) \\ \tau s' &= -s.\end{aligned}\tag{9}$$

The additional variable s describes the exponentially-decaying transient input. At tone onset, this variable is instantaneously displaced to $s(t_{on}) = [I_T - \beta I_N]_+$. Similar dynamics occur at tone offset, with γ_{on} replaced by γ_{off} and $s(t_{off}) = -[I_T - \beta I_N]_+$. Notice $s(t_{off})$ is negative-valued because offset responses are inhibitory inputs to x .

We configured Model 2 so that it is bistable in the absence of any inputs. In the x - s phase plane, bistability comprises stable equilibria at inactive and active firing rates (x_I and x_A , respectively) separated by an unstable saddle point (x_S). All equilibria are located along the s -axis. Activation of the population from rest requires that the transient onset response is sufficiently large to transition x from the basin of attraction of x_I to the basin of attraction of x_A . This condition can be visualized in the phase plane by considering the separatrix curve $S(x)$ that divides these two basins of attraction. The firing rate will activate from rest if $s(t_{on}) > S(x_I)$, that is if the response variable at tone onset exceeds the height of the separatrix curve evaluated at the inactive firing rate equilibrium. We observed that the separatrix curve can be adequately approximated as a line connecting the saddle point ($x_S, 0$) to the point ($x_I, 1$). The choice of $s = 1$ at threshold enforces our convention that activation for tone-only inputs occurs at $I_T = 1$. From this geometric argument, the linear approximation to the separatrix is

$$S(x) = \frac{x_S - x}{x_S - x_I}\tag{10}$$

Trajectories in the phase plane at tone onset are shown in Fig. 4A1 and their full time-courses are shown in Fig. 4A2. If the tone level is sufficiently high ($I_T = 1.2$ in this simulation, red curve), the system crosses the separatrix and transitions to the stable equilibrium in the active state.

This approximation to the separatrix can also be found by linearizing the dynamical system in Eq. 9 about the saddle point and determining the eigenvector associated with its stable manifold. The Jacobian matrix for the system is

$$J(x, s) = \begin{bmatrix} (-1 + a_e f'(a_E x + \gamma_{on} s)) / \tau & \gamma_{on} f'(a_E x + \gamma_{on} s) / \tau \\ 0 & -1 / \tau \end{bmatrix} \quad (11)$$

297 The negative eigenvalue for the saddle point is the lower right entry of this matrix. The associated
298 eigenvector satisfies $(J_{11} - J_{22})x + J_{12}s = 0$, where J_{ij} is the (i, j) entry of the Jacobian matrix so we
299 conclude that the linear approximation to the separatrix is

$$S(x) = \left(\frac{J_{22} - J_{11}}{J_{12}} \right) (x - x_S). \quad (12)$$

300 A useful consequence of our assumption that x and s have the same time constant τ is that this expression
301 can be simplified substantially to:

$$S(x) = \frac{a_E}{\gamma_{on}} (x_S - x). \quad (13)$$

302 Comparing this equation to the result obtained by geometric considerations (Eq. 10), we observe that γ_{on}
303 is determined by the parameters a_E and m , and our convention that tone threshold is $I_T = 1$. In particular,
304 we set $\gamma_{on} = a_E(x_S - x_I)$.

305 Deactivation is the mirror image of activation and occurs if the “downward” perturbation of s is sufficiently
306 strong at the tone offset. For the values of a_E and m that we use, the stable equilibria are symmetric around
307 the saddle point at $x = 0.5$ and it is convenient to set $\gamma_{on} = \gamma_{off}$ so that the threshold for activation and
308 deactivation are the same. More generally, the offset parameter should always be set so that activation
309 thresholds are not less than deactivation thresholds, to avoid the scenarios in which a tone onset can activate
310 the firing rate variable but the tone offset response is too weak to return the population. In this unrealistic
311 case, x could remain in the active state for perpetuity.

312 3.3.2 Response dynamics to tone masked by noise

313 If a sufficiently strong noise is presented at the same time as a tone, then the noise can prevent activation of
314 the tone population by reducing the transient response at the start of the tone. This is the masking condition.
315 Simulations exhibiting masking dynamics are in Fig. 4B. Recall the effect of noise is to reduce the onset
316 response to $s(t_{on}) = [I_T - \beta I_N]_+$ where β is a parameter that controls how much noise suppresses the
317 tone onset (given in Eq. 4). As we explained above, in our discussion of tone activation, the model is
318 parameterized so that $I_T = 1$ is the threshold for activation in the absence of noise. Thus, a noise will mask
319 a tone if $I_T - \beta I_N < 1$. We denote the threshold for masking as $M(I_T)$ and conclude that it is related to
320 tone strength via a linear equation:

$$M(I_T) = \frac{1}{\beta} (I_T - 1), \quad (14)$$

321 This equation is valid for values of I_T above the noise-free threshold ($I_T = 1$) and below the maximum
322 tone strength in the model ($I_{T,max} = 5$). It provides a direct relationship between masking threshold and

the degree to which noise masks tone onsets (represented by the parameter β). In the simulations shown in Fig. 4) we use $\beta = 2/3$. The masking threshold curve is shown in Fig. 4D.

3.3.3 Response dynamics to tones interrupted by noise

The feature of this model that is essential for our study of the continuity illusion is that the inactive and active coexist and are stable in the absence of any inputs. The tone population can, therefore, remain in the active state even after a tone is turned off if the offset signal is weak and does not send x across the separatrix. Simulations exhibiting masking dynamics are in Fig. 4C.

Whereas masking depends on suppression of the onset response (as described above), the continuity illusion depends on suppression of the offset response. Computation of the continuity threshold is analogous to our derivation of the masking threshold, but with onset terms replaced by offset terms. The criteria for continuity are that $I_T > 1$ (so that the first tone activates the population) and that I_N is sufficiently large to reduce the tone offset response and allow the x firing rate to remain near the upper equilibrium state. To satisfy this second condition, we must have that the offset response does not cross the separatrix curve, with the key difference being that we are now analyzing the system at the start of the noise-filled gap. This means $x(t_{off}) = x_A$ (first tone has activated the population) and $s(t_{off}) = -[I_T - \beta I_N]_+$.

Adapting the separatrix equation in Eq. 13 for the offset response, we have that continuity requires $-[I_T - \beta I_N]_+ < a_E(x_S - x_A)/\gamma_{off}$. The continuity threshold equation is, therefore,

$$C(I_T) = \frac{1}{\beta} \left(I_T - \frac{a_E}{\gamma_{off}}(x_S - x_A) \right). \quad (15)$$

In the particular case of Model 2, we have chosen parameters that make it symmetric (x_I and x_A are equidistant from the saddle point). We set $\gamma_{off} = \gamma_{on} = a_E(x_S - x_I)$, so that the continuity and masking thresholds are identical (compare $C(I_T)$ to Eq. 14. Masking and continuity threshold lines are identical and plotted in Fig. 4D).

More generally, we would require $\gamma_{off} \geq \gamma_{on}$ to avoid persistent activation (discussed above). If $\gamma_{off} > \gamma_{on}$, the continuity threshold would shift upward in Fig. 4D (γ_{off} affects the intercept of $C(I_T)$ but not its slope). In the intermediate I_N values between $C(I_T)$ and $M(I_T)$, we observe responses to interrupted tones in which the first tone activates x , the noise does not cause firing to persist (no continuity) and the noise prevents activation by the second tone. In other words, the model without symmetry results in a region in stimulus space in which the noise burst between the two tones is too weak to induce the continuity illusion but sufficiently strong to prevent perception of the second tone by forward masking.

3.4 Model 3: Combined inputs implement both rules for the continuity illusion

The last model configuration we consider is one that cannot be activated by transient inputs alone or sustained inputs alone. These requirements are met if I_T at the left knee of the I_T - x equilibrium curve is positive (no bistability at rest) and the I_T at the right knee is to the right of the maximum allowable tone level $I_{T,max}$. Activation from rest can only occur using a combination of sustained and transient inputs. There must be a sufficiently strong sustained input to move the system past the saddle node bifurcation point at the left knee of the equilibrium curve. This creates a stable equilibrium in the activated state that can be accessed if the transient portion of the input is sufficiently strong to transition the system into the basin of attraction of this upper equilibrium, see Fig. 5.

Figure 5. Activation dynamics in Model 3 requiring sustained and transient inputs. (A) Firing rate equilibrium curve. Defining feature of Model 3 is a left knee at positive I_T level and right knee at I_T level larger than $I_{T,max}$. (B) Time-courses of firing rate variable for sustained input only (blue), transient input only (red), and combination of both inputs (yellow). Tone level is $I_T = 1.5$ in all cases.

To understanding firing rate responses for this model, we again formulate the dynamics in the x - s state space:

$$\begin{aligned}\tau x' &= -x + f(a_E x + [I_T - \beta I_N]_+ - a_I(1-x)I_N + \gamma s) \\ \tau s' &= -s\end{aligned}\tag{16}$$

These equations govern the dynamics of the system in the time following a tone onset or offset. The parameter γ should be thought of representing γ_{on} to describe onset responses or γ_{off} for offset responses. The initial values for these equations are given by the state of the systems immediately prior to sound onset or offset. For the case of tone onset for a system starting from rest (no input), for instance, the initial values would be $x(t_{on}) = x_I(0, 0)$ and $s(t_{on}) = 1$, where $x_I(0, 0)$ is the inactive state in the case of $I_T = I_N = 0$. We will use similar notation throughout this section to indicate that equilibrium points are functions of the input levels.

The Jacobian matrix for these equations is

$$J(x, s) = \begin{bmatrix} (-1 + (a_e + a_I I_N) f'(u)) / \tau & \gamma f'(u) / \tau \\ 0 & -1 / \tau \end{bmatrix}\tag{17}$$

where we have abbreviated the argument of f' with $u = a_E x + [I_T - \beta I_N]_+ - a_I(1-x)I_N + \gamma_{on}s$. As before, we evaluate the Jacobian at the saddle point (when it exists, for sufficiently large I_T) and use the eigenvector associated with the stable manifold of the saddle point to construct a linear approximation to the separatrix curve that defines the threshold for activation. The notable difference between the analysis in this section and the preceding section (for Model 2), is that the sustained input terms can affect the positions of the equilibrium solutions and the shape of the separatrix in the current model setting. Following our earlier calculation (recall Eq. 12), we find the eigenvector by solving $(J_{11} - J_{22})x + J_{12}s = 0$. After simplifications, we find the linear approximation to the separatrix to be

$$S(x; I_T, I_N) = \left(\frac{a_E + a_I I_N}{\gamma} \right) (x_S(I_T, I_N) - x)\tag{18}$$

where we are assuming that I_T is sufficiently large so that the saddle point $x_S(I_T, I_N)$ exists. In the remaining sections, we apply this result in the three cases we have been considering (tone only, simultaneous tone and noise, and tones with a noise-filled gap) to characterize activation by tone, masking, and continuity dynamics.

3.4.1 Response dynamics to tone-only inputs

Activation by a tone-only input ($I_N = 0$) occurs if the onset response causes the system to cross the separatrix defined in Eq. 18. In particular, we consider the system starting from rest, with $x(t_{on}) = x_I(0, 0)$ and input variable s instantaneously perturbed to $s(t_{on}) = I_T$. We then ask if this onset perturbation to s exceeds $S(x(t_{on}))$. We must also keep in mind that the position of the saddle point is determined by inputs, so in this case we use $x_S(I_T, 0)$ in Eq. 18. From these considerations we conclude that, to satisfy our convention that $I_T = 1$ is the tone threshold, we must set the onset parameter to $\gamma_{on} = a_E (x_S(1, 0) - x_I(0, 0))$. Simulations showing activation by a tone are shown in Fig. 6A. The phase

390 portrait in Fig. 6A1 illustrates the dynamics at tone onset. We remark that I_{offset} is not necessary to move
 391 the system back to the inactive state. In this model, the return to the inactive state at the end of the tone is
 392 guaranteed in the tone-only case because the saddle point and upper equilibrium do not exist for $I_T = 0$.
 393 The firing rate variable must return to $x_I(0, 0)$ because it is the unique, remaining stable equilibrium. This
 394 differs from Model 2 which required an offset response to deactivate x . We will see shortly, however, that
 395 the offset response does affect dynamics of the continuity illusion dynamics, and we will explore γ_{off}
 396 further in that setting.

Figure 6. Dynamics of Model 3 responses to tone and noise inputs. (A) Tone inputs (no noise) with tone level $I_T = 1.2$. Phase plane at tone onset, showing stable equilibria (black dots), unstable saddle (gray circle), linear approximation to separatrix (gray line) and trajectory (blue curve). Time-course of firing rate variable $x(t)$ shown in panel A2. (B) Simultaneous tone and noise inputs (masking condition) with $I_T = 2$ and $I_N = 1.5$. Phase plane shown at tone and noise onset. The active state still exists, but it is shifted rightward relative to tone only and slope of the separatrix is steeper. Noise prevents activation in this case and firing rate returns to inactive equilibrium. (C) Interrupted tone with noise-filled gap (continuity condition) with $I_T = 2$ and $I_N = 1.5$. Phase plane shown at noise onset. Noise preserves stable equilibrium. In this case, trajectory does not cross separatrix and system remains active during gap between tones. (D) Masking threshold and continuity thresholds. Parameter values used: $a_I = 7$, $\alpha = 0.5$, $\beta = 0.05$, $\gamma_{on} = 9.6$, $\gamma_{off} = 0.88$. (E) Continuity threshold for varying γ_{off} parameter value. Black curve in (D) and (E) are identical.

397 3.4.2 Response dynamics to tone masked by noise

398 In the masking condition, tone and noise inputs are both present simultaneously. As a result, the saddle
 399 point about which we linearize the system depends on the inputs $x_S(I_T, I_N)$. To determine whether a noise
 400 prevents activation by a tone, we are still interested in onset dynamics from rest. The initial firing rate is
 401 $x(t_{on}) = x_I(0, 0)$. We use the separatrix equation (Eq. 18) again to determine the masking threshold. We
 402 determined γ_{on} and we use it to update the separatrix at tone onset:

$$S(x; I_T, I_N) = \left(1 + \frac{a_I}{a_E} I_N\right) \left(\frac{x_S(I_T, I_N) - x}{x_S(1, 0) - x_I(0, 0)}\right). \quad (19)$$

403 This equation reveals the differing effects of sustained tone and noise inputs. The sustained tone input does
404 not alter the slope of the linear approximation to the separatrix, it is the same as the slope found for the
405 model with transient inputs only (Eq. 12), but it can change the intercept through the dependence of x_S on
406 I_T in the numerator. The sustained noise input, in contrast, changes the intercept and also the slope of the
407 intercept (through the term $a_I I_N / a_E$).

408 Trajectories in the phase plane and as firing rate time-courses showing masking of a tone by noise are in
409 Fig. 6B. In this example, the onset response exceeds the threshold for tone-only inputs ($s(t_{on})$ is above one
410 in Fig. 6B1). Nevertheless, the population returns to the inactive state because the noise input has raised
411 the threshold for activation (in this case, primarily by steepening the slope of separatrix).

412 With Eq. 19 in hand, we compute the masking threshold as the minimum I_N level (for a given I_T level)
413 at which the onset response crosses the separatrix. To do this, we set the amplitude of the transient onset
414 response to the height of the separatrix at the initial value. Thus, the masking threshold $M(I_T)$ is the
415 noise level that solves the nonlinear equation $I_T - \beta I_N = S(x_I(0, 0); I_T, I_N)$. We solve this equation
416 numerically and display the resulting masking threshold curve in Fig. 6D, for selected parameter values.

417 3.4.3 Response dynamics to tones interrupted by noise

418 In contrast to the two cases just considered (activation by a tone alone, masking by noise), understanding
419 the continuity illusion requires analysis of the offset response. We approximate the threshold, as usual,
420 with the linear approximation to the separatrix that is given in Eq. 18. We interpret γ in that equation as
421 γ_{off} since we are concerned with the dynamics at the offset of the first portion of the tone. There is no
422 tone ($I_T = 0$) during the noise-filled interruption between the tones, so the saddle point position is only a
423 function of I_N . This highlights the fact that there are two necessary conditions for continuity dynamics
424 in this model: the noise level must be sufficiently strong to preserve the saddle point and stable active
425 equilibrium and the offset response must be sufficiently weak (perhaps because it is masked by the noise)
426 to prevent the system from crossing the separatrix. The second of these conditions is satisfied if the offset
427 response, which has initial amplitude $-(I_T - \beta I_N)$, does not cross $S(x_A; 0, I_N)$. A firing rate response
428 that remains active when the tone is removed (consistent with the continuity illusion) is shown in the x - s
429 phase plane in Fig. 6C1 and as a time-course in Fig. 6C2.

430 The offset parameter γ_{off} shifts the system between the two extreme cases represented by Models 1 and
431 2. If the offset parameter is small, then the occurrence of the continuity illusions relies on the hysteresis
432 effect (the fact that I_N can preserve the saddle point and active equilibrium). In this case, the continuity
433 threshold $C(I_T)$ changes slowly with I_T (and, in fact, can be constant over a range of I_T values). In the
434 extreme case of no offset response, the continuity threshold would be constant with I_T as was the case for
435 Model 1 (see Fig. 3D). If the offset parameter is large, then the offset response makes a larger contribution
436 to whether the firing persists during the gap. The continuity threshold varies more with I_T as γ_{off} increases.
437 These effects of γ_{off} on continuity threshold are shown in Fig. 4D.

438 4 DISCUSSION

438 The continuity illusion is an intriguing example of the capacity for the brain to “fill in” missing information.
439 In this case, this “filling in” process creates an illusion of a continuous tone that is, in fact, discontinuous.
440 In the context of hearing in a complex listening environment, a bias toward linking related sounds across

time to create longer-lasting auditory objects can be useful when contending with multiple interrupting or distracting sounds that could momentarily obscure a sound of interest. Our contribution has been to use principles of neural dynamics and to draw on previous conceptual frameworks and experimental studies of the continuity illusion to identify dynamical mechanisms that can account for the continuity illusion. Our approach was to use a firing rate model as a caricature of population-level neural activity. Firing rate models have been used in studies of visual [17, 18, 19] and auditory illusions [20, 21, 22] and are useful for identifying possible dynamical and neural mechanisms that underlie perceptual outcomes.

We explored the possible routes toward continuity illusion-like dynamics. A first requirement was that recurrent excitation be strong enough to create a stable equilibrium at a high firing rate level (Fig. 2). We stereotyped the inputs to the population as sustained, transient onset, and transient offset. This setup was based on physiological evidence that neurons in auditory cortex with these response types are possible neural correlates for the continuity illusion [8]. Although the model could be configured so that sustained inputs alone (Model 1) or transient inputs alone (Model 2) produce dynamics consistent with the continuity illusion, there are shortcomings in both cases that can be remedied in a model that requires both sustained and transient inputs together to activate the population (Model 3). In the case of sustained inputs alone, continuity dynamics require a hysteresis effect: input levels that are too weak to activate the population from rest can, nevertheless, sustain activity in an already active population. The tone input makes no contribution to the response dynamics during the noise-filled interruption between the tones. As a result, the continuity threshold is constant with tone level (Fig. 3D). This is inconsistent with evidence that the probability of perceiving the continuity illusion increases as the noise level becomes louder relative to the tone [23]. In the case of transient inputs alone, continuity dynamics require a population that is bistable in the absence of any inputs. If additional mechanisms were not at work (synaptic adaption in the recurrent excitatory connections, for instance), a bistable population could remain in a high firing rate state for a long period of time even when no stimulus is present. The model that requires both inputs to activate (Model 3) can utilize both input types to implement the continuity illusion. Sustained noise-driven inputs are necessary to preserve a stable equilibrium in the active state (via a hysteresis effect) and transient offset responses can ensure that the continuity threshold is not independent of tone level (Fig. 6H).

Following Petkov et al. [8], we view the different input populations (sustained, onset, and offset) as possible neural correlates of two aspects of the descriptive theory of the continuity illusion proposed by Bregman [1]. First, sustained responses convey “evidence” that a tone is ongoing. If some sustained neural activity persists during the interruption between two tones, then the tone may be perceived as continuous. This is Bregman’s “Sufficiency of Evidence rule.” In our model, this is implemented by assuming that noise (as a broadband signal) drives a partial excitatory input to the population whose activity signals the perception of the tone. Second, transient responses at tone onsets and offsets mark “discontinuities” in the tone (acoustic edges). If the noise during the gap between tones obscures this discontinuity, then tone may be perceived as continuous. This is Bregman’s “No Discontinuity rule.” In our model, this is implemented by assuming noise can reduce the amplitude of the transient response. An additional effect of noise (in Model 3) is to increase the threshold for activation by a transient input (effected by altering the separatrix, see I_N term in Eq. 19).

By analyzing how these input types drive firing rate activity, we identified the dynamical mechanisms and circuit properties that can support the continuity illusion. A hysteresis effect enabled the noise to convey “sufficient evidence” to sustain firing activity in the gap between tones. A suppression of the offset response due to noise kept removed the “discontinuity” in the acoustic signal and confined the system to the basin of attraction of the high firing rate stable equilibrium. These dynamical mechanisms (hysteresis

and bistability) required sufficiently strong recurrent excitation, specifically the requirement that $a_E > 4$ for S-shaped equilibrium curves (Fig. 2).

In addition to providing a dynamical framework to accompany Bregman's theory, our approach offers a new perspective on previous models of the continuity illusion. The work of Noto et al. [11] relied on nonlinear, self-excitation of neural populations which is consistent with our approach for creating hysteresis and bistable dynamics. The work of Vinnik et al. [12] utilized dynamic synapses to create continuous responses to interrupted tones. Since dynamic synapses are an additional mechanism for creating hysteresis and bistable dynamics [24], there is also consistency between their work and ours. An extra feature of our study, not considered in these previous models of the continuity illusion, is that we also required that our simulated dynamics exhibit masking for noise and tones presented simultaneously. Incorporating this is essential, in our view, to avoid "trivial" solutions in which noise acts simply (and solely) as an excitatory input that facilitates neural activity during the noise-filled gap between tones. Balancing these excitatory and inhibitory effects of noise required fine-tuning in Model 1 (Fig. 3E).

Although the firing rate framework is a caricatured description of neural activity, insights into a number of sensory illusions have been made using similar approaches including for visual bistability [17, 18, 19] and auditory streaming [20, 21, 22]. In our study, it has been useful as a minimal model that identifies mechanisms that can account for the continuity illusion. We focused on a classic version of the continuity illusion (tone interrupted by noise), but the illusion can be elicited by a variety of sound types [1, 2, 3]. The facts that the continuity illusion is widespread and that the dynamics of the illusion can be explained with a relatively simple model may indicate that circuitry that promotes perceptual continuity is present at multiple auditory centers. Indeed, while we were motivated by observations in auditory cortex [8], the input motifs we used in our model (sustained, onset, and offset responders) are present in multiple auditory nuclei. See, for instance, Rhode and Smith [25] for sustained and onset response types in the ventral cochlear nucleus and Kopp-Scheinpflug et al. [13] for descriptions of offset neurons at multiple nuclei in the auditory system). There is also evidence for involvement of the brainstem in the continuity illusion in human listeners [9]. These observations may support a view that perceptual continuity of interrupted sounds is constructed at more than one stage of auditory processing. As work continues to identify neural correlates of the continuity illusion, idealized dynamical systems descriptions can inform what features should be included in future, more biophysically-detailed modeling approaches.

RESOURCE IDENTIFICATION INITIATIVE

Computations were performed using MATLAB (RRID:SCR_001622), version R2018b.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

JHG conceived of the study. All authors designed, implemented, and carried out simulations and analysis. All authors contributed to writing and editing the manuscript.

FUNDING

This research has been supported by Swarthmore College through the Deborah A. DeMott '70 Student Research and Internship fund (Qianyi Cao) and the Eugene M. Lang Summer Research Fellowship (Noah Parks).

DATA AVAILABILITY STATEMENT

Matlab code for the firing rate model and to generate all figures in this study can be found in the github repository at <https://github.com/jhgoldwyn/ContinuityIllusion>.

REFERENCES

- [1] Bregman A. *Auditory scene analysis: The perceptual organization of sound* (MIT Press) (1990).
- [2] Warren RM. *Perceptual restoration of missing sounds* (Cambridge University Press), chap. 6. 3 edn. (2008), 150–173. doi:10.1017/CBO9780511754777.007.
- [3] McWalter R, McDermott JH. Illusory sound texture reveals multi-second statistical completion in auditory scene analysis. *Nature Communications* **10** (2019) 5096. doi:10.1038/s41467-019-12893-0.
- [4] Miller G, JCR L. The intelligibility of interrupted speech. *Journal of the Acoustical Society of America* **22** (1950) 167. doi:10.1121/1.1906584.
- [5] Warren RM. Perceptual restoration of missing speech sounds. *Science* **167** (1970) 392–393. doi:10.1126/science.167.3917.392.
- [6] Husain FT, Lozito TP, Ulloa A, Horwitz B. Investigating the neural basis of the auditory continuity illusion. *Journal of Cognitive Neuroscience* **17** (2005) 1275–1292. doi:10.1162/0898929055002472.
- [7] King AJ. Auditory neuroscience: Filling in the gaps. *Current Biology* **17** (2007) R799–R801. doi:10.1016/j.cub.2007.07.013.
- [8] Petkov CI, O'Connor KN, Sutter ML. Encoding of illusory continuity in primary auditory cortex. *Neuron* **54** (2007) 153–165. doi:10.1016/j.neuron.2007.02.031.
- [9] Bidelman GM, Patro C. Auditory perceptual restoration and illusory continuity correlates in the human brainstem. *Brain Research* **1646** (2016) 84–90. doi:10.1016/j.brainres.2016.05.050.
- [10] Riecke L, Micheyl C, Oxenham AJ. Global not local masker features govern the auditory continuity illusion. *Journal of Neuroscience* **32** (2012) 4660–4664. doi:10.1523/JNEUROSCI.6261-11.2012.
- [11] Noto M, Nishikawa J, Tateno T. An analysis of nonlinear dynamics underlying neural activity related to auditory induction in the rat auditory cortex. *Neuroscience* **318** (2016) 58–83. doi:10.1016/j.neuroscience.2015.12.060.
- [12] Vinnik E, Itskov P, Balaban E. A proposed neural mechanism underlying auditory continuity illusions. *J Acoust Soc Am.* **128** (2010) EL20–5. doi:10.1121/1.3443568.
- [13] Kopp-Scheinflug C, Sinclair JL, Linden JF. When sound stops: Offset responses in the auditory system. *Trends in Neurosciences* **41** (2018) 712–728. doi:10.1016/j.tins.2018.08.009. Special Issue: Time in the Brain.
- [14] Miller P. *An Introductory Course in Computational Neuroscience* (MIT Press), chap. 6 (2018).
- [15] Brody CD, Romo R, Kepecs A. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology* **13** (2003) 204–211. doi:10.1016/S0959-4388(03)00050-3.
- [16] Warren RM, Obusek CJ, Ackroff JM. Auditory induction: Perceptual synthesis of absent sounds. *Science* **176** (1972) 1149–1151. doi:10.1126/science.176.4039.1149.
- [17] Laing CR, Chow CC. A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience* **12** (2002) 39–53. doi:10.1023/A:1014942129705.
- [18] Shpiro A, Curtu R, Rinzel J, Rubin N. Dynamical characteristics common to neuronal competition models. *Journal of Neurophysiology* **97** (2007) 462–473. doi:10.1152/jn.00604.2006.
- [19] Shpiro A, Moreno-Bote R, Rubin N, Rinzel J. Balance between noise and adaptation in competition models of perceptual bistability. *Journal of computational neuroscience* **27** (2009) 37. doi:10.1007/s10827-008-0125-3.

- 564 [20] Rankin J, Sussman E, Rinzel J. Neuromechanistic model of auditory bistability. *PLOS Computational*
565 *Biology* **11** (2015) 1–34. doi:10.1371/journal.pcbi.1004555.
- 566 [21] Rankin J, Osborn Popp PJ, Rinzel J. Stimulus pauses and perturbations differentially delay or promote
567 the segregation of auditory objects: Psychoacoustics and modeling. *Frontiers in Neuroscience* **11**
568 (2017) 198. doi:10.3389/fnins.2017.00198.
- 569 [22] Paredes-Gallardo A, Dau T, Marozeau J. Auditory stream segregation can be modeled by neural
570 competition in cochlear implant listeners. *Frontiers in Computational Neuroscience* **13** (2019) 42.
571 doi:10.3389/fncom.2019.00042.
- 572 [23] Riecke L, van Opstal AJ, Formisano E. The auditory continuity illusion: A parametric investigation
573 and filter model. *Perception & Psychophysics* **70** (2008) 1–12. doi:10.3758/PP.70.1.1.
- 574 [24] Barak O, Tsodyks M. Persistent activity in neural networks with dynamic synapses. *PLOS*
575 *Computational Biology* **3** (2007) 1. doi:10.1371/journal.pcbi.0030035.
- 576 [25] Rhode WS, Smith PH. Encoding timing and intensity in the ventral cochlear nucleus of the cat.
577 *Journal of Neurophysiology* **56** (1986) 261–286. doi:10.1152/jn.1986.56.2.261.