# Locating phonological explanation in the neural dynamics of speech: hysteresis and coupling

Jason A. Shaw

*Yale University*

## 1 Mentalism in phonology

Phonologists may disagree about whether phonology is cognitive science or not. Formal phonology, like other areas of grammar, can be understood as characterizing a speaker/hearer's *knowledge of language*. On this view, a 'mentalist' perspective, phonology is an aspect of human cognition, and models of phonology are models of cognition, albeit at a certain level of abstraction. Alternatively, the very same formal models of phonology can alternatively be understood from a strictly 'empiricist' perspective, as characterizing patterns in language in a rigorous and insightful manner, without commitment to cognition or, more specifically, the cognitive processes that generate language. On this perspective, the phonological grammar provides a model of phonological data. Capturing the systematicities and idiosyncrasies with precision while bringing insight to the question of possible phonological patterns is a central goal of the field, regardless of whether we take a 'mentalist' or 'empiricist' perspective. That similar formal tools have been used by practicing phonologists from both perspectives may be a sign that commitment to mentalism (or not) has so far been largely orthogonal to the core work of building formal models of phonology.

In this paper, we pursue a formal approach to phonology that effectively disambiguates 'mentalist' and 'empiricist' perspectives on grammar. We present neural process models of phonological patterns, showing how phonological alternations derive from the neural dynamics that support cognition. We'll see that the explanation for common phonological patterns derive from basic principles of dynamical systems, hysteresis and coupling, as applied to neural activation. On this perspective, phonological typology relates directly to parameterization of neural dynamics.

## 2 Data

**2.1** *Symbolic nature of phonological data* Phonological data is, by convention, typically represented as sequences of symbols, often drawn from the International Phonetic Alphabet (IPA), representing categories of sound. The data themselves, therefore, are represented at a symbolic level of abstraction from measureable signals, e.g., neural activation, muscle contraction, articulatory movement, acoustics, etc. There are known limitations to such representations, particularly when it comes to equating across languages (Pierrehumbert et al. 2000). For example, the same IPA symbol may have different phonological status across languages (e..g., Mielke 2005; Sande and Oakley 2023). Nevertheless, abstracting speech to sequences of symbols has been insightful, supporting identification of patterns that are now core data for phonological theory. In keeping with this convention, we present phonological data in terms of IPA symbols and interpret the output of the dynamical models in the same way, even as the explanation for the patterns comes from the neural and articulatory dynamics. The empirical focus will be synchronic patterns of dissimilation, as exemplified in (2.2), and assimilation, as exemplified in (2.3).

* I'd like to thank the organizers of AMP 2025 and the audience for their feedback as well as Shigeto Kawahara for discussion in early developments of the research and the Yale Dynamics group, especially Michael Stern, Manasvi Chaturvedi, Finn Amber, Alessandra Pintado-Urbanc, Rina Furusawa, and Anton Hampe, who provided comments on a draft of the paper. Remaining errors are entirely my responsibility. For Matlab code to run the simulations reported in the paper, please contact me by email: jason.shaw@yale.edu


Proceedings of *AMP 2025*

**2.2** *Dahl's law in Kikuyu* In a typological survey, Suzuki (1998) identifies instances of phonological dissimilation on nearly every dimension of phonological contrast, including major place, lateral, rhotic, voice, spread glottis, constricted glottis, nasal, continuant, vowel features (high, low, back), length, and tone (high, low). For the sake of empirical concreteness, we situate our discussion of dissimilation in the context of a well-studied voicing dissimilation rule, known as Dahl's law, with specific reference to Kikuyu, also known as Gikuyu, a stable Eastern Bantu language of Kenya (per Ethnologue). Davy & Nurse (1982) argue convincingly for the synchronic status of one form of the rule. The general pattern, which has traces in various forms across Eastern Bantu, is for the first of two consecutive syllable-initial voiceless stops to become voiced. In Kikuyu, specifically, the active synchronic process only targets velar stops with the trigger being any voiceless stop in the following syllable onset. The synchronic alternation can be seen in prefixes, beginning with voiceless velar stops, when they are added to stems that also begin in voiceless stops. Examples with the diminutive prefix, /kɑ-/, are provided in (1).

(1) Prefix /kɑ-/ used as a diminutive (data based on Davy & Nurse 1982)[1]

| | | | |
|---|---|---|---|
| [goko] 'chicken' | [kɑ-goko] 'small chicken' | [ko] 'firewood' | [gɑ-ko] 'small firewood' |
| [βori] 'goat' | [kɑ-βori] 'small goat' | [çerɑ] 'path' | [gɑ-çerɑ] 'small path' |
| [ɦiti] 'hyena' | [kɑ-ɦiti] 'small hyena' | [tɛgwɑ] 'ox' | [gɑ-tɛgwɑ] 'small ox' |

**2.3** *Inter-vocalic voicing* As an example of phonological assimilation, we consider inter-vocalic voicing, the pattern whereby an underlying voiceless consonant is voiced between vowels. In some languages, there is evidence from alternations that this pattern has a phonological status. The data in (2) come from Sekani, a Na-Dene language of Canada (Hargus 2018). The forms in (a), (b), (c) are nouns that begin with voiceless fricatives. The forms in (d), (e), (f) have a vowel-final prefix before the stem, which puts the underlying voiceless consonant in stem-initial position between two vowels. In this environment, the consonant surfaces as voiced.

(2) Vowel-final prefix causes stem-initial voiceless fricatives to voice. (Hargus 2018: 229-230)

| | |
|---|---|
| a. [xas] 'planning tool' | d. [sə-ɣase] 'my planning tool' |
| b. [çen] 'song' | e. [sə-ʝen] 'my song' |
| c. [xaz] 'windfall roots' | f. [tse-ɣaz] 'Old Friend Mt.' |

Intervocalic voicing is one of the more commonly observed lenition patterns (Gurevich 2013). Even in languages in which it has not been phonologized, it can be observed probabilistically—voiceless stops variably surface as voiced in numerous (unrelated) languages—e.g., Rome Italian & Basque (Nadeu and Hualde 2015), Tohoku Japanese (Mizoguchi et al. 2020), and Sanuma (Borgman 1990). Even languages not described as having this pattern—such as English—show some degree of gradient intervocalic voicing of stops (Davidson 2016). This is a phonological pattern with a clear phonetic precursor.

## 3 Dynamical systems

**3.1** *Dynamical systems as models of reality* Dynamical systems are the standard mathematical tools for describing the natural world. A dynamical system describes change in the state of a system over time. The basic setup includes a state variable (or set of variables), which define a state space, and a rule/law (or a set of rules/laws), the "dynamic", which specifies how the state variable changes over time. The formal expression of a dynamical system is a differential equation (or system of differential equations); for a general overview, see Strogatz (2024); for a speech-specific introduction, see Iskarous (2017)

When considering state variables, measurement is an important consideration. It may be difficult or impossible to measure the state variable of maximal theoretical interest. In general, we may have to cede that we do not have perfect (or even good) measurements of some of the most theoretically interesting state variables for language; nevertheless, we hope to have measurements that are at least a lawful function of

[1] Since the emphasis on this paper is just voicing, I transcribe the voiced velar as [ɡ], which differs from Davy & Nurse (1982), who transcribe the voiced velar as [ɣ], noting that it typically has the character of a "voiced, weakly fricative or frictionless approximant".

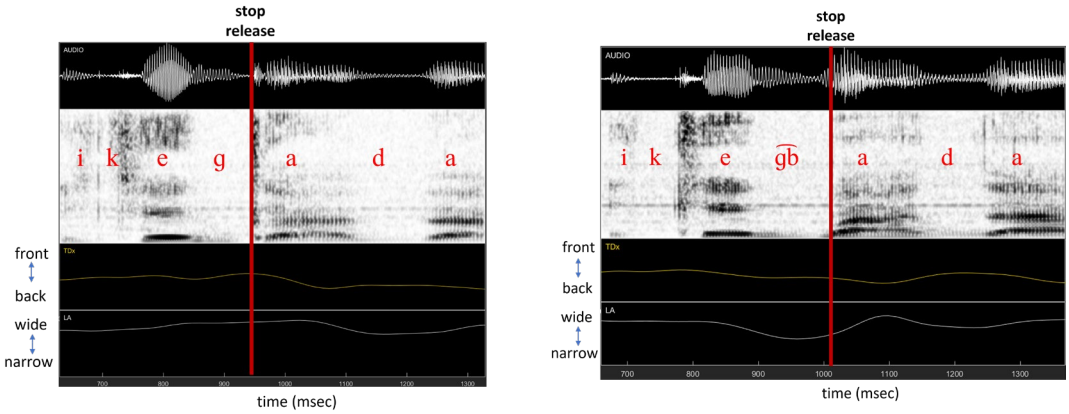
variables of interest. Research applying dynamical systems approaches to languages have considered a range of state spaces, which differ in their measurability. For example, Nowak et al. (2002) set up Universal Grammar (UG) as their state variable and theorize its dynamics. Although this is a state variable of theoretical interest to linguists, it is exceedingly difficult (and maybe impossible) to measure. To model stress patterns, researchers have set up dynamical systems with syllable prominence as the state variable (Goldsmith and Larson 1990; Iskarous and Goldstein 2018), which may be easier to approximate by measurement. There are also applications of dynamical systems to diachronic change; for example, the state variable set up by Sonderegger and Niyogi (2013) is the 'probability of a rule applying', a variable that can be estimated by conducting the same phonological analysis at different diachronic stages of a language. There has also been a wave of research modelling the neural activation underlying language production as the state space (Chaturvedi and Shaw 2025a; Chaturvedi and Shaw 2025b; Gafos and Kirov 2010; Kirkham and Strycharczuk 2024a; Kirov and Gafos 2007; Pintado-Urbanc 2025; Roon and Gafos 2016; Shaw 2025; Stern 2025; Stern et al. 2022; Stern and Shaw 2023), which this paper will build on.

**3.2** *Dynamical systems models of articulation* The above examples not withstanding, a much more common language-relevant state space has been goal-driven articulatory movement (e.g., Browman and Goldstein 1989; Browman and Goldstein 1990; Byrd and Choi 2010; Elie et al. 2023; Katsika et al. 2014; Kelso et al. 1986; Kirkham 2025; Nam et al. 2009; Saltzman and Munhall 1989; Stern and Shaw 2025). Measurement of this state variable has improved substantially over the last several decades in tandem with the rise of Laboratory Phonology and interest in relating phonological structure to the continuous variables of articulation. Articulatory movements are attractive state variables for dynamical systems in part because of their measurability.

An example of articulatory trajectories recorded with Electromagnetic Articulography (EMA) is provided in (3). The figure shows acoustic and articulatory trajectories as displayed in MVIEW, Matlab-based software for data visualization (Tiede 2005). Trajectories show minimal pairs /ga/ and /g͡ba/ in Dangme (Kwa language of Ghana).[2] The top panel shows the acoustic waveform. Below the waveform, the figures show a spectrogram and two articulatory trajectories: (i) the longitudinal (front to back) position of the tongue dorsum (TDx), shown in red and (ii) lip aperture (LA), the Euclidean distance between the upper and lower lips, shown in green. A vertical red line marks the stop release.

(3) Articulatory trajectories comparing minimal pairs /ga/ and /g͡ba/ in Dangme (Kwa language of Ghana).

/ike#**ga**#da/ /ike#**g͡ba**#da/




A clear difference between the examples is that the labial-velar /g͡ba/ (right) shows a marked decrease in lip aperture, bottom panel, before the stop release, indicating the lips are closed, while the plain velar /ga/ (left) shows a relatively wide lip aperture during the stop closure. From this data, we can conclude that there is a

[2] Data in (1) comes from Beauty Korkor Amamor, a native Dangme speaker, recorded at the Bureau of Ghana Languages in Accra, Ghana in December 2024; thank you to BGL staff Enoch Adibuer, Rachel Agyeman, David Kloteye, and Phoebe Welbeck for assistance with the experiment.

controlled movement of the lips for /g͡ba/ that is absent in /ga/. Additionally, there are clear differences in the longitudinal position of the tongue dorsum (TDx). The tongue dorsum is more retracted (back) in /g͡ba/ than in /ga/. The symbolic representation of the minimal pairs, /ga/ vs. /g͡ba/, anticipates the presence of a labial closure in /g͡ba/, which is absent in /ga/. However, it does not represent the difference in tongue dorsum retraction.

In contrast to symbolic representations, dynamical representations readily capture both the difference in lip aperture as well as the difference in tongue dorsum retraction shown in (1). Pioneering research in this area—the framework of Articulatory Phonology—modelled phonological patterns with dynamical systems, setting up goal-directed movement as the state space for the dynamical system (e.g., Browman and Goldstein 1986; Browman and Goldstein 1990). The dynamics adopted in early work was the critically damped mass spring system. Subsequent research has sought to refine the dynamics, improving empirical fit to data and theoretical interpretability (e.g., Byrd and Saltzman 1998; Elie et al. 2023; Kirkham 2025; Sorensen and Gafos 2016; Stern and Shaw 2025).

There are, by now, several competing proposals for the dynamics of controlled articulatory movement, with a rich literature discussing empirical and theoretical arguments in favor of one or another. These proposals differ in details, which have important implications, but they also have a lot in common. The papers cited above all propose systems that can be characterized in dynamical terms as point attractors; the systems evolve over time to a particular state, which serves as equilibrium for the system. Conceptually, the point attractor in these systems is the target of the movement or the goal of speech production. Thus, two phonological categories are differentiated if they have different articulatory goals, i.e., different equilibrium points in state space. To highlight their similarities and differences, (4) arranges the formal expression of dynamical proposals (as differential equations) solved for their highest derivative, with the parameter dictating the location of the point attractor expressed as T, for target.

(4) summary of some of the proposed dynamics for phonologically relevant speech movement control

| | **Dynamical system** | **Representative Citation** | **Notes** |
|---|---|---|---|
| (a) | $\ddot{x} = k(-x + T) - 2\dot{x}\sqrt{k}$ | Browman & Goldstein 1989 | Critically damped mass spring |
| (b) | $a_t\ddot{x} = a_t k(-x + a_t\mathrm{T}) - 2\dot{x}$ | Byrd & Saltzman 1998 | (a) with addition of activation ramp |
| (c) | $\ddot{x} = k(-x + \mathrm{T}) - 2\dot{x}\sqrt{k} + d(x - \mathrm{T})^3$ | Sorenson & Gafos 2016 | (a) with non-linear restoring force |
| (d) | $\dot{x} = -2(-x + T)/k(t - d^2/t)$ | Elie et al., 2023 | Based on General Tau theory |
| (e) | $\ddot{x} = r\dot{x} - \dot{x}^2/(-x + \mathrm{T})$ | Stern & Shaw 2025 | Derived from simple point attractor |
| (f) | $\ddot{x} = -bx + k(-x + T)$ | Kirkham 2025 | Discovered with machine learning |

The important point for the remainder of this paper is that there is broad agreement from these proposals—although they differ in important ways—that there is a parameter, T, which dictates the direction in which the state space will evolve. Returning to our example in (3), we can express the difference between /ga/ and /g͡ba/ in Dangme in terms of the presence/absence of controlled movement for the lips together with a different value of the T parameter for the tongue body constriction location: more posterior for /g͡ba/ than for /ga/.

**3.3** *Connecting articulation to phonology* When compared to state spaces such as "Universal Grammar" or even "Syllable Prominence", the prospect of relating the state space of "goal-directed movement" to measurements of the physical world is relatively tractable. Indeed, some of the theoretical development represented in (4) has come by holding our theories accountable for an increasing degree of precision in fitting model parameters to measurements of articulatory data (Elie et al. 2023; Kirkham 2025; Stern and Shaw 2025). As satisfying as it is to explain substantial variance in empirical measurements, the state space of goal-directed movement is not the state space of most interest to phonology. The observations that are the core data of the field, expressed in terms of symbols, show that the same linguistic entity, a lexical item, can vary in its phonological exponent across contexts. Certain of these patterns, phonological alternations, may lawfully derive from the overlap of gestures, with no contextual change in parameters, as discussed in the Articulatory Phonology literature (e.g., Browman and Goldstein 1989; Goldstein 2011; Parrell and Narayanan 2018). However, other phonological patterns do require contextual change in parameters (e.g., Ellis and Hardcastle 2002; Kochetov and Pouplier 2008; Shaw and Kawahara 2023). In

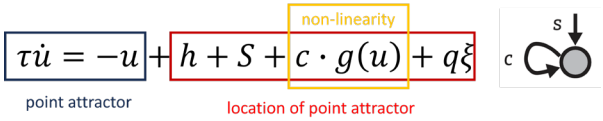
relation to the articulatory dynamics described above, accounting for phonological patterns requires modelling systematic changes in the T parameter across contexts. That is, we need a dynamics for T, a parameter dynamics, which captures contextual influences. Since we understand T as dictating the target of phonologically relevant movement, we turn to theories developed in cognitive neuroscience for human action control, building our T parameter dynamics from the neural dynamics of action control.

## 4 The neural dynamics of nodes and fields

The key components of our neural dynamic architecture, which we will use to generate the patterns in (1) and (2), are neural nodes and neural fields, which we introduce in this section. The formal definitions follow Schöner and Spencer (2016). Neural fields have insightfully modelled the influence of competitors and context on the selection of movement parameters in speech (Burroni 2023; Chaturvedi and Shaw 2025a; Chaturvedi and Shaw 2025b; Kim and Tilsen 2025; Kirkham and Strycharczuk 2024b; Roon and Gafos 2016; Shaw 2025; Stern et al. 2022; Tilsen 2019), proposals which generalize aspects of the neural dynamics developed for other human actions, e.g., such as eye and hand movements (Erlhagen and Schöner 2002). Neural nodes have played less of a role in models of action and perception (outside of speech), but they may be crucial to phonology.

**4.1** *Neural node dynamics* We assume that phonological computation is a neural process and adopt neural activation as the state variable for our dynamical model. Neural activation, represented as the variable $u$, corresponds to the average spike rate across a population of neurons. Change in neural activation is represented as $\dot{u}$. Change is function of current state, $u$, resting level, $h$, input, $S$, self-excitation, $c \cdot g(u)$, and noise, $q\xi$. The rate of change in activation is modulated by $\tau$. The formal expression of neural node dynamics is given in (5)

(5) Formal expression of neural node dynamics, following Schöner and Spencer (2016)




Like the various proposals for gesture dynamics, neural nodes have point attractor dynamics. This qualitative description of the dynamics derives from the first part of the equation in (5), labeled "point attractor". Change in activation is a negative function of current activation. The location of the point attractor is determined by the other terms in the equation, which are added together. By convention, we set $h$ to -5, which means that the system will evolve to -5 when all other terms are 0. We can think of a node as being in either an "off" state or an "on" state. By convention, 0, is the threshold between states, so that negative values of the activation variable indicate the node is "off" while positive indicate it is "on". Thus, when the input $S$ is greater than the resting value, $h$, the point attractor will be positive, driving the system towards the on state. Self-excitation provides a boost, adding activation—shifting the point attractor further in a positive direction—just when activation crosses the initial threshold. This functions to stabilize the activation of neurons, preventing them from quickly falling back into the off state. The self-excitation term, $c \cdot g(u)$, introduces non-linearity into the dynamics. The function $g(u)$ is a sigmoid, returning values close to 0, when $u$ is negative, and values close to 1, when $u$ is positive. Thus, an activation "boost" of strength $c$ kicks in, just when $u$ is greater than 0.

**4.2** *Neural node coupling* The same formal mechanism for self-excitation, described for a single node above, also models coupling between nodes. The coupling dynamics are given in (6) for two nodes, which each have activation values, $u_1$, for node1 and, $u_2$, for node2. Both nodes have resting levels, inputs and self-excitation, as in (5). The additional term is coupling from $u_1$ to $u_2$. When $u_1$ is in the on state (> 0), then $g(u1)$ will approximate 1, and $c_{21}$, the strength of the coupling on node2 from node1 will be added to the node2 activation. In the examples we consider, coupling strength, $c_{21}$, between a lexical node and a phonological node is excitatory (positive values of $c_{21}$) while coupling between phonological nodes is inhibitory (negative values of $c_{21}$), which means that an active node1 drives the point attractor of node2

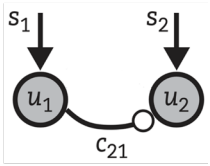
towards the off state.

(6) Coupled node dynamics following Schöner and Spencer (2016)

$$\tau\dot{u}_1 = -u_1 + h + \boldsymbol{S}_1 + c_{11} \cdot g(u_1)$$

point attractor — location of point attractor — coupling term

$$\tau\dot{u}_2 = -u_2 + h + \boldsymbol{S}_2 + c_{22} \cdot g(u_2) + \boldsymbol{c_{21} \cdot g(u_1)}$$




**4.3** *Dynamic Neural Fields* A Dynamic Neural Field (DNF) is a population of neurons that is sensitive to a particular metric dimension. Each location in a DNF is sensitive to a different value of the relevant metric dimension. DNFs are used for both perception and action, and, indeed, have been used to model the interaction between perception and action in speech (Roon and Gafos 2016). Our focus here is production.

Since our aim is to capture laryngeal patterns, in (1) and (2), the metric dimension of interest is Glottal Width (GW). Our DNF for GW is 100 neurons wide. Each location in the field is preferentially sensitive to a different value of GW. The field ranges from one, or closed glottis, through narrow glottis values, the state for voicing, to wide glottis, the state for voicelessness. We have expressed the units of GW as a percentage of maximum GW so that the range goes from 1 to 100. We assume that GW of less than 40% is the voicing state while a GW of greater than 40% is the voiceless state.

As we are modeling action control with this DNF, a peak at some location in the field corresponds to an intended action. For example, a peak at 60 in the DNF corresponds to an intension to move towards 60% open glottis. Within our broader architecture, a peak within the DNF supplies the target, T parameter, to the gesture dynamics. The peak in the field does not represent the physical state of GW but rather the intention for action in a particular direction. We can therefore understand the DNF in the context of speech action as representing an 'intentional dynamics'.

The neural node dynamics introduced in (5) and (6), including self-excitation and inhibition, generalize to fields with some modest modifications. The formal expression of a DNF is provided in (7).

(7) Dynamics Neural Field, following Schöner and Spencer (2016)

Change of activation — Resting activation

$$\tau\dot{u}(x,t) = -u(x,t) + h + s(x,t) + \int k(x-x')g\big(u(x',t)\big)dx' + q\xi(x,t)$$

activation — input — Interaction kernel (field-based equivalent of excitatory and inhibitory neural interaction) — noise

$$s(x,t) = a\exp\left[-\frac{(x-p)^2}{2w^2}\right]$$

The state space of the DNF is neural activation, $u$, the same as nodes, with the addition that there is activation at each location in the field, $x$. The change in activation, $\dot{u}$, over time, $t$, at each location in the field, $x$, is represented by $\dot{u}(x,t)$. This is a function of current activation at each location in the field at a given time, $-u(x,t)$, as well as a resting level, $h$, inputs, $s(x,t)$, and noise, $q\xi(x,t)$. The inputs to the field take the form of Gaussian distributions, with three parameters, input location, $p$, input width, $w$, and input amplitude, $a$. Finally, there are lateral interactions within the field (coupling between field locations), represented in compact form by $k(x-x')$, which includes three components, (1) local excitation, (2) local inhibition, and (3) and global inhibition. Global inhibition applies uniformly across the field but local excitation and local inhibition have Gaussian shapes, with amplitude, position, and width parameters. These lateral interactions are gated by a sigmoid, $g\big(u(x',t)\big)$, so they only kick in once some location in the field crosses threshold, similarly to the node interactions already described. In practice, the amplitude of local excitation is set to greater than the sum of local inhibition and global inhibition. Thus, when one location of the field crosses threshold, it receives a boost, i.e., self-excitation. Neighboring neurons, those sensitive to similar degrees of

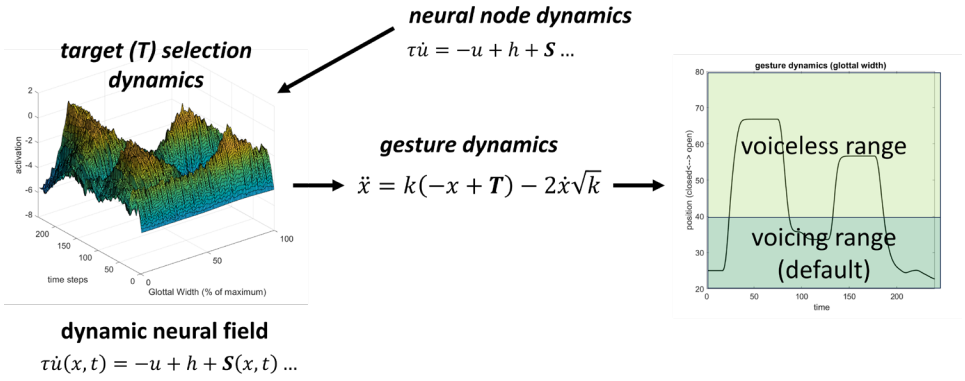
GW, also receive a boost, owing to the Gaussian shape of local excitation, while neurons that are more distal in the field, those outside the scope of local excitation, will be inhibited, due to the global inhibition component, and possibly the local excitation component, which is, by convention, set to have a wider width (and lower amplitude) than local excitation.

The dynamics of neural fields make interesting predictions for action sequences requiring multiple peaks in the same DNF (see also Chaturvedi and Shaw 2025a; Shaw 2025). The interaction kernal functions to effectively stabilize an action target against noise, allowing peaks in the field to be maintained long enough to be successfully executed (by supplying the target for movement). This function is crucial for action behavior. Without self-excitation, the DNF may flutter above and below threshold causing pathological starts and stops to movement; without global inhibition, the location of the activation peak may shift within the DNF, changing the target before it can be achieved. Interestingly, self-excitation and global inhibition, which are fundamental to stable action behavior, for reasons described above, have consequences that can be recognized as phonological patterns.

## 5 Simulations

A schematic diagram of the neural process architecture is provided in (8). It shows gesture dynamics (right), represented by the damped mass-spring equation with parameter T, coupled to a DNF which selects the value of T, which is coupled to neural nodes, which provide the input parameters to the DNF. Using the architecture in (8), with dynamic nodes providing input to fields which select parameter values for gesture dynamics, we simulate the phonological patterns in (1) and (2) and highlight the dynamical source of phonological alternation for each pattern. The role of ***coupling*** between dynamic nodes is crucial for dissimilation, (1), while ***hysteresis*** within the DNF accounts for assimilation, (2).

(8) Schematic depiction of three-level (nodes, fields, gestures) dynamic architecture. The T parameter of the gesture dynamics is selected from a dynamic neural field, "target selection dynamics", which receives input from neural nodes, "neural node dynamics". For concreteness, we consider the dimension of "glottal width", where narrow glottal width conditions voicing and wide glottal width conditions voicelessness.



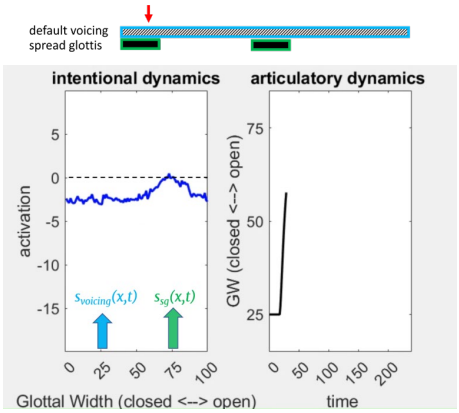

**5.1** *Assimilation as hysteresis* We first apply the architecture to assimilation, which emerges from field dynamics. Because of global inhibition, a component of the DNF interaction kernal, an activation peak sustained in one location of the field will suppress activation in other locations in the field *below resting level*. Consider an input to our DNF of sufficient amplitude and duration to drive activation from resting level activation to across-threshold activation, causing a peak to form in the DNF. If the same input is applied to the field in a location that has been inhibited below resting level, it can fail to induce an activation peak. This is a case of hysteresis in that the state of the system is dependent on its state history. Inhibiting one portion of the DNF can prevent a subsequent peak from forming in the location of inhibition. Instead, the DNF is biased by self-excitation to maintain activation peaks in stable locations.

This general mechanism of neural processing for action control offers a natural explanation for progressive assimilation of the type exemplified in (2). That is, an input to the voiceless end of the GW field of sufficient duration and amplitude to cause a peak when the field is at rest, as in the stem-initial voiceless

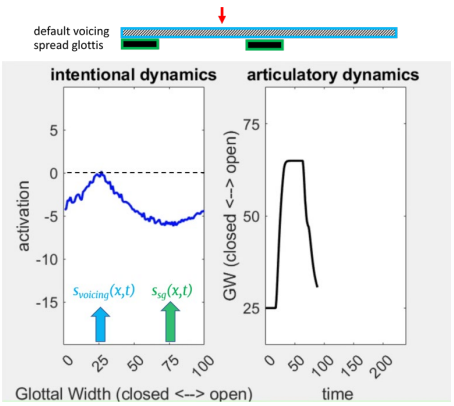
cases (2a), (2b), (2c), is not sufficient to cause a peak following global inhibition from a sustain peak at the voiced end of the field. Thus, when following a vowel, which corresponds to a period of voicing induced by a peak in the DNF at the voicing end of the field, the input to the voiceless end of the field fails to induce a peak. This leads to voicing through the inter-vocalic consonant closure.

(9) Simulation: evolution of the intentional dynamics (GW DNF) and articulatory dynamics (GW gesture) at four time steps. A schematic representation of the timing of inputs to the field is given at the top of each panel with the red arrow indicating the timeslice represented in DNF activation (left) and GW state (right).

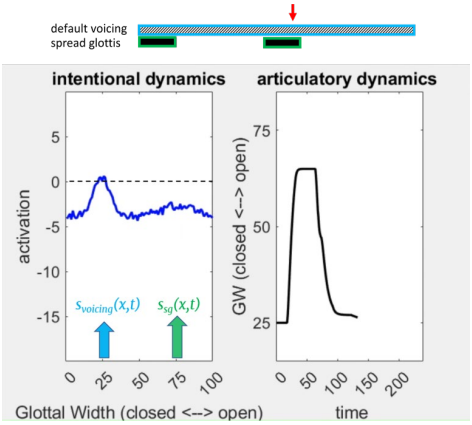
(time 1: voiceless GW state)



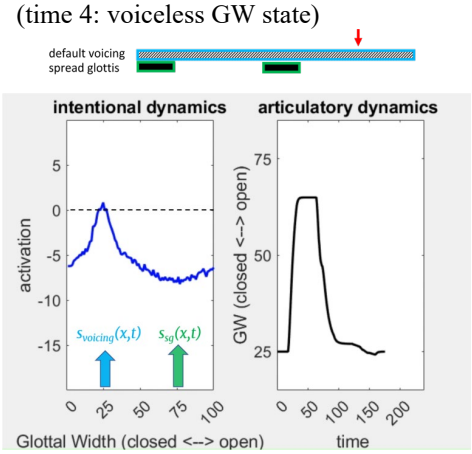

(time 2: voiced GW state)




(time 3: voiced GW state)




(time 4: voiceless GW state)




The simulation in (9) illustrates intervocalic voicing, showing the state of the intentional dynamics (left), represented by the GW DNF, and the gesture dynamics (right) at four time points. The resting level of the field is -5 ($h$ = -5) and we assume that voicing is the default speech-ready state, which we model with an ever-present input at the narrow GW end of the field ($p$ = 25) of amplitude 5 (a = 5), just strong enough to drive the field to threshold (in the absence of other inputs). We overlap the default voicing input with a spread glottis input, i.e., at the wide GW end of the field; this is the input associated with underlyingly voiceless obstruents, such as the fricatives in (2). Our simulation includes two such voiceless inputs spaced over time, one for the first voiceless stop and one for the second. This first input occurs at the beginning of the simulation; time 1 in (9). The second occurs much later, at time 3, after a long period with only the default voicing input, time 2. Both spread glottis inputs have the exact same parameters. They are centered on 75 ($p$ = 75) with an amplitude of 7 ($a$ = 5). Note that the amplitude of the spread glottis inputs is higher than the default voicing input ($a$ = 5). This means that—all else equal— the neurons at the spread glottis end of the field will evolve towards the "on state" (above zero) faster than the neurons at the voicing end of the field. Once activation at the voiceless end of the field crosses threshold, it will inhibit the voicing end of the field. When the field starts at resting level uniformly, as it does at the start of this simulation, we can predict that

the activation peak will emerge at the spread glottis (voiceless) end of the field, due to the higher amplitude of the spread glottis input than of the default voicing input. We see this at time 1 (9: top, left)—there is an activation peak at the spread glottis end with the voicing end of the field remaining above rest level (due to default voicing input) but below threshold (due to global inhibition from the spread glottis peak). At time 2 (top, right), a peak has formed at the voicing end of the field (left), under the influence of default voicing. The spread glottis end of the field is inhibited below resting level, due to global inhibition and the absence of any counteracting input. The gesture dynamics move towards the voiced state (right).

Continuing description of the simulation in (9), the second spread glottis input enters the field, at time 3 (bottom, left). This time, the spread glottis input follows an epoch (time 2) in which a robust voicing peak had been established. The spread glottis end of the field is therefore inhibited to the point that, even the stronger amplitude of the spread glottis input is insufficient to drive activation across threshold in the time that it is active. At time 3, spread glottis input increases activation at the wide GW end of the field, but because it is starting from a below resting level state, an activation peak does not ultimately form. Thus, despite the spread glottis input, the GW state remains in the voicing range. When the spread glottis input is removed, at time 4 (bottom right), then global inhibition from the voicing peak pushes the high end of the GW back below resting level, while the GW state remains in the voicing range.

To summarize, assimilation in the case we have described is categorical at the level of the gesture dynamics. That is, the gesture dynamics never receive a voiceless target, even though there is an active input to the DNF at the wide (spread glottis) end of the field. The reason is that the immediately preceding activation peak inhibited the spread glottis end of the field below threshold. While intervocalic voicing is relatively common, it is not universal. Some languages show consistent realization of voiceless stops as voiceless between vowels. We can understand such variation in terms of the dynamic parameters. For example, Shaw (2025) shows that scaling the duration of the spread glottis inputs—holding all other parameters constant—is sufficient to derive three distinct patterns: intervocalic voicing (as we have seen), voicing contour, and vowel devoicing. Increasing the duration of the spread glottis inputs will allow them enough time to overcome inhibition (from default voicing). However, if they are too long, then they can make it less likely for default voicing to cross threshold, leading to vowel devoicing. Thus, the same type of hysteresis that causes voicing assimilation, as in (9), can cause devoicing, under different DNF inputs. More broadly, then, we can understand typological variation in phonological patterns in terms of the parameters of the dynamical system that governs speech production.

### **5.2** *Dissimilation from node coupling*

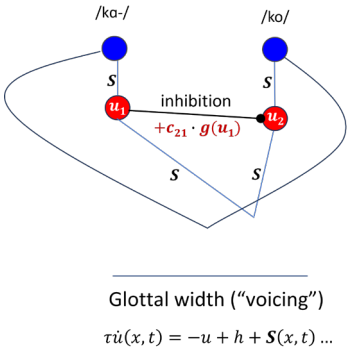
While the key explanatory mechanism in the assimilation model was DNF hysteresis, this mechanism alone cannot account for dissimilation. Accounting for dissimilation requires interaction at the level of neural nodes, which provide input to the field. We model both lexical items and phonological features as neural nodes. A schema of the architecture, provided in (10), shows a total of four nodes—two for lexical items (blue) and two for spread glottis features (red). It also shows one DNF, representing glottal width (GW). The phonological feature nodes are coupled to the high end (right side) of the field. When activation of these nodes crosses threshold, they send excitatory input to the high (large glottal width) end of the field. A peak in this end of the field ultimately conditions voicelessness by driving the articulatory dynamics to a wide glottis state. The phonological nodes (red) receive input, S, from lexical nodes (blue), just when the activation of the lexical nodes crosses threshold. In our simulations, the lexical nodes receive input externally and then excite the rest of the system. The lexical nodes are also coupled directly to the glottal width field, but they provide broad input to the middle of the field. Direct coupling from lexical nodes to fields enables the possibility of word-specific phonetics that are nonetheless also constrained by sub-lexical representations (Shaw and Tang 2023). We do not make use of this potential here, as we lack the relevant data. Nevertheless, the coupling from lexical nodes to the DNF raises the resting level of the field when a lexical item is active, encouraging default voicing. We assume that there is a weak default input to the small glottal width ("voicing") end of the field, as we did for the assimilation simulation in (9). This will lead to an activation peak at the "voicing" end of the glottal width field, unless active input at the voiceless end is stronger.

Considering Dahl's law, (1), the lexical nodes in (10) represent items such as the diminutive prefix /kɑ-/ and stems that condition voicelessness, such as /ko/. The key difference between the left and right side of (10) is the direction of coupling between the phonological nodes. On the left, node1 ($u_1$) inhibits node2 ($u_2$);

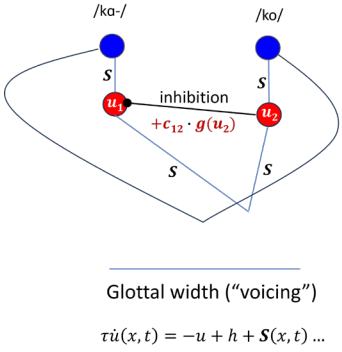
on the right, the directionality is reversed, node2 ($u_2$) inhibits node1 ($u_1$).

(10) Parametric variation in node coupling directionality

(a) Node1 ($u_1$) inhibits Node2 ($u_2$)



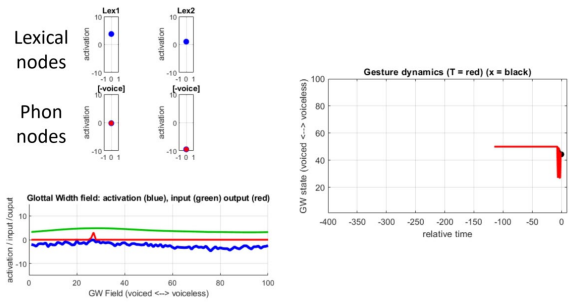

(b) Node2 ($u_2$) inhibits Node1 ($u_1$)



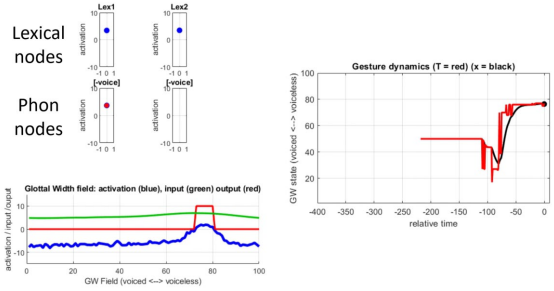

The directionality of node-to-node coupling has interesting phonological consequences. We first consider the architecture in (10a), where Node1 ($u_1$) inhibits Node2 ($u_2$). This coupling relation enforces serial order. We illustrate this through simulation in (11), which shows changes to node activation (top left), field state (bottom, left), and gesture dynamics (right) at four time slices.

(11) Simulation: serial order derives from Node1 ($u_1$) inhibiting Node2 ($u_2$)

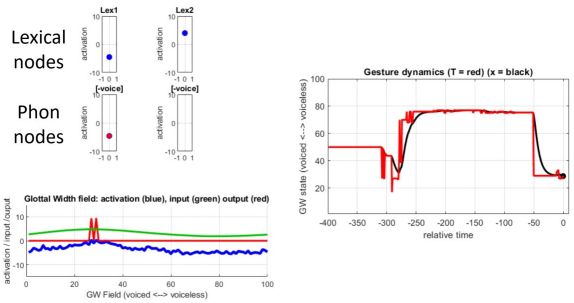
(time 1: starting GW state)



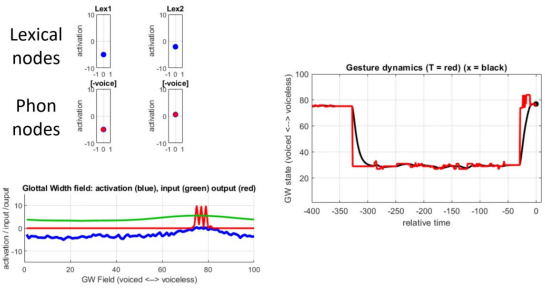

(time 2: voiceless GW state)




(time 3: voiced GW state)




(time 4: voiceless GW state)




At time 1 (top, left), both lexical items have received input. The input is greater for lexical item1 (lex1) than lexical item 2 (lex2), so lex1 is well across threshold while lex2 is only slightly positive. The input from the above-threshold lex1 to phon node1 ([-voice] feature for Lex1) drives phon node1 across threshold, which, inhibits phon node 2, driving it lower (greater negative activation) into the off state. Also at time 1, the glottal width (GW) field has a weak input at the voicing end of the field (from default voicing) and the lexical boost. The gesture dynamics are in a neutral state but trending slightly towards voicing, because of the target provided by default voicing. Note that the gesture dynamics panel shows the T parameter (in red), which is the location of the activation peak in the DNF. This changes overtime as the field evolves, but the parameter is "smoothed" by the gesture dynamics, which move continuously toward the point attractor.

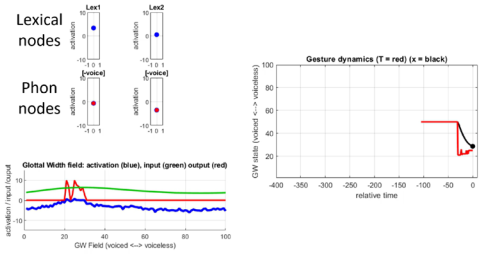
At time 2, both lex nodes and phon node1 are above threshold, phon node2 has been completely inhibited (off of the scale), and a peak has formed at the voiceless (wide glottal width) end of the field, which has driven the gesture dynamics in that direction. At time 3, lex node1 has returned to rest position, as has phone node 1. Lex node2 is still above threshold but phon node2 is still recovering from inhibition (activation is too low to show up on the scale). While phon node2 recovers from inhibition, a peak in the field forms at the voicing end, under the influence of the default voicing input and lexical boost from node2. At time 4 (bottom, right), phon node2 recovers from inhibition, crosses threshold and causes a peak at the voiceless end of the field, driving gesture dynamics back in this direction.

To summarize, over the course of the simulation, this dynamical system provides a sequence of voiceless, voiced, voiceless glottal states. Because node1 inhibits node2, it allows for a period of voicing to intervene between the voiceless specifications. In the absence of this inhibition, both nodes would exert voiceless influence on the field at the same time. The temporal interval between the two voiceless influences, with voicing in-between, emerges from Node1 ($u_1$) inhibiting Node2 ($u_2$).

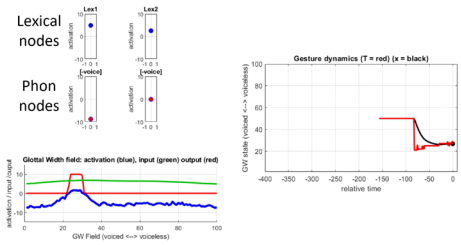
We next examine what happens when the coupling goes in the other direction, i.e., when node2 ($u_2$) inhibits node1 ($u_1$). The simulation results are presented in (12). The panels follow the same organization as in (4). At time 1, both lexical nodes have crossed threshold, which provides a boost to the whole GW field allowing default voicing to cross threshold and drive the gesture dynamics into the "voicing" range for glottal width, before either of the phon nodes cross threshold. At time 2, phon node2 crosses threshold and inhibits phon node1 before it can induce a peak in the GW field. The gesture dynamics continue towards the voicing target and remain in that state. At time 3, lexical node1 has fallen below threshold and phon node 1 remains well-below threshold, due to inhibition from phon node 2. Phon node2 remains above threshold, providing input to the voiceless end of the GW field. A peak in that part of the field supplies the gesture dynamics with a voiceless target, which drives GW out of the voicing range. A time 4, the gesture dynamics achieve the GW target provided by the field while phon node1 remains inhibited well below resting activation. Thus, in this simulation, the first lexical item (lex node1) activates a [-voice] feature (phon node1), but before that voiceless feature can supply a speech production target, it is inhibited by another voiceless feature (phon node 2), which is excited by the second lexical item (lex node2).

(12) Simulation: regressive dissimilation from Node2 ($u_2$) inhibiting Node1 ($u_1$)

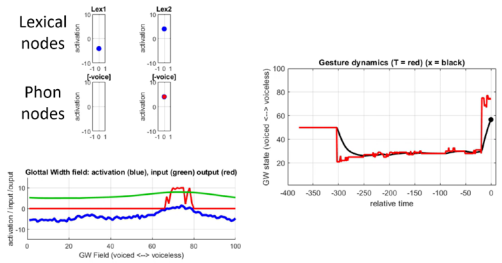
(time 1: neutral GW state)



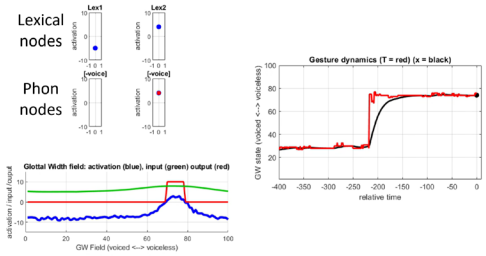

(time 2: voiceless GW state)




(time 3: voiced GW state)




(time 4: voiceless GW state)




Ultimately, phon node2 does cause a peak to form at the wide (voiceless) end of the GW field; however, this takes time and glottal wide remains in the "voiced" range for a significant time interval—enough, on this

account (about 250 ms, interpreting each time step as a millisecond), to produce the first syllable as voiced.

Simultaneous activation of two lexical items, modelled as neural nodes, enables phonological features, also modelled as nodes, to interact. Depending on the direction of the coupling between the nodes, two behaviors are possible: imposition of a serial order in how the features impact speech production (11) or regressive dissimilation (12), whereby one feature "loses" its opportunity to influence speech production. This occurs because, during the temporal window when the lexical item activates the feature node, it is being inhibited by a simultaneously active node. When the inhibition ends (because node2 falls below threshold), the source of input to phon node1 (the lex node1) has also fallen below threshold.

**5.3** *Discussion of simulations* The intuition captured by the model is that simultaneously active neural representations interact to yield what can be identified as categorical outcomes in speech production. Continuous variation in glottal width results in discontinuities in voicing state, i.e., quantal transitions between "voiced" and "voiceless" states (Stevens 1989). In (11), the glottal state goes from "voiceless" to "voiced" and back to "voiceless". In (12), the glottal state starts "voiced" and remains "voiced" until a final transition to "voiceless". These categorical jumps between glottal states are the result, in the model, of a gesture dynamics that drives glottal state continuously over time towards a target, which is "selected" from a DNF that is both spatially and temporally continuous. Activation peaks in the DNF, which evolve over time under the input of nodes, determine the gesture target.

The nodes that provide input to the DNF do not have a metric dimension—they are simple activation variables (dynamic nodes). In isolation they have no phonetic content, but they are coupled to a particular location of the continuous DNF, namely, the voiceless (high glottal width) end, which is why we label them [-voice]. Even as they have no phonetic content themselves, activation of these nodes can impact speech production—driving the system to a voiceless state—because of how they are coupled to the DNF and, additionally, how they are coupled to other nodes.

The precedence (serial order) of voiced/voiceless in the gesture dynamics derives entirely from activation strength and coupling. Of the lexical nodes, we have labelled one "Lex1" and the other "Lex2". Lex1 crosses threshold before Lex2 because it has higher activation strength. Serial order (precedence) at the level of the lexical nodes is thus determined by input activation strength, a long-established method in neural networks (e.g., Jordan 1985). In the simulations, input to Lex1 and Lex2 is supplied at the same time, but the input amplitude is stronger in Lex1 than in Lex2, causing Lex1 to cross threshold first. The coupling between lexical and phonological feature nodes is the same for Lex1 and Lex2. Thus, all else equal, it would follow that the [-voice] node coupled to Lex1 would cross threshold before the [-voice] node coupled to Lex2. What disrupts this outcome is the coupling between the [-voice] nodes. In (12), the [-voice] node excited by Lex1 fails to cross threshold, because it is first inhibited by the [-voice] node excited by [Lex2]. In the absence of this inhibition, the first [-voice] node would cause a peak to form—this is what we see in (11). Thus, the context of lex2 conditions the absence of a voiceless target during the first lexical item. This serves as our model of dissimulation of the type observed in Dahl's law, (1), where the presence of a voiceless stop conditions the realization of "preceding" voiceless stop as voiced.

In the simulations reported in (11) and (12), the value of the $\tau$ parameter, controlling the timecourse of the dynamics, differed by "layer". The lexical nodes had a higher $\tau$ value than the phonological nodes. This means that the activation variable for the lexical nodes changed more slowly than the phonological nodes. The intuition behind this parameter setting is that sub-lexical nodes must have a faster time course than lexical nodes so that the rise and fall in activation of each sub-lexical node can be triggered within the time of lexical node. In this way, the $\tau$ parameter facilitates encoding a kind of hierarchical structure in that faster evolving nodes (phonological) complete their activation cycles within the temporal life of the slower evolving nodes (lexical). As a principle, superordinate neural populations evolve at a slower timescale than subordinate neural populations. More broadly, this parameter setting is consistent with the observation that human auditory system naturally tracks multiple timescales, the faster gamma band (31-45 Hz) for gestures and the slower theta band (4-7 Hz) for syllables (Teng et al. 2017). There is also some psychological reality to the different timescales from speech error research—segmental errors tend to be local, occurring within a few syllables, while lexical errors have a much wider temporal scope, often occurring several syllables away (Fromkin 1971).

Coupling between nodes is gated by a sigmoid function, which dictates that interaction—one node impacting another—only happens when that node crosses threshold, advancing to the "on state". Weak input

to a node may be insufficient to drive it across threshold (unless combined with input from multiple sources) (c.f., Smolensky and Goldrick 2016). Because of self-excitation, nodes that do cross threshold are likely to stay in the "on state" long enough to exert some influence on the activation of other nodes or a field, although this depends as well on hysteresis, i.e., the state of the field and the state the of the gesture dynamics. For example, if the GW gesture is already in a voiceless state, then an activation peak at the voiceless end of the GW field may not cause articulatory movement (because the gesture dynamics are already at target/equilibrium). Similarly, if the GW field has a stabilized activation peak at the voiced end of the field, input at the voiceless end will have less impact (and potentially no measurable impact) than if activation was at rest. These consequences follow from properties of the DNF and their integration in the larger dynamical system.

## 4 Conclusion

The case studies described in this paper show how two common phonological patterns, dissimilation and assimilation, can be modelled in terms of the neural process that generates them. Progressive assimilation has a natural explanation in DNF dynamics. In contrast, regressive dissimilation of the Dahl's Law variety, cannot be derived from the duration and amplitude of inputs into a DNF. This requires interactions between neural nodes, which, themselves, are simple activation variables (no phonetic content). The cascade of activation needed to derive dissimilation also included nodes for lexical items, which necessarily evolve at a slower timescale than phonological feature nodes. Assimilation and dissimilation, thus, on this neural process account, have qualitatively different dynamics.

We simulated neural processes using the tools of Dynamic Field Theory (DFT) to represent the intentional dynamics, i.e, the selection of targets for gestures. The DFT models were implemented with the COSIVINA toolbox (Schneegans 2021) in Matlab, augmented with new elements to interface with gesture dynamics and to facilitate coupling between nodes and fields. The dynamical models make predictions that are in principle testable against a range of measurements, including brain data and articulatory measurements. Our focus here was on a categorical interpretation of the gesture dynamics in term of voicing states, which maps directly to IPA symbols classified as voiceless, e.g., [k], [x], [ç], or voiced, e.g., [g], [ɣ], [ʝ]. Although symbols offer only a coarse level of description, which may not be sufficient for comparing across languages (Pierrehumbert et al. 2000), they offer a starting point for honing in on a language-specific parameterization of neural dynamics. For cognitive neuroscience, they offer a useful and hitherto under-utilized source of data constraining our models of neural dynamics. For phonology, neural process models offer new potential to understand attested (and possible) patterns in terms of the cognitive capacity that gives rise to them.

## 5 References

Borgman, Donald M. 1990. Sanuma: Mouton de Gruyter Berlin.

Browman, Catherine & Louis Goldstein. 1986. Towards an Articulatory Phonology. Phonology Yearbook 3.219-52.

—. 1989. Articulatory gestures as phonological units. Phonology 6.201-51.

—. 1990. Gestural Specification Using Dynamically-Defined Articulatory Structures. Journal of Phonetics 18.299-320.

Burroni, Francesco. 2023. Dynamics of F0 planning and production: Contextual and rate effects on Thai tone gestures: Cornell University Ph. D. Dissertation.

Byrd, Dani & Susie Choi. 2010. At the juncture of prosody, phonology, and phonetics: the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. Laboratory Phonology, ed. by C. Fougeron, B. Kühnert, M. D'Imperio & N. Vallée, 31-59. Berlin, New York: Mouton de Gruyter.

Byrd, Dani & E. Saltzman. 1998. Intragestural dynamics of multiple prosodic boundaries. Journal of Phonetics 26.173-99.

Chaturvedi, Manasvi & Jason A Shaw. 2025a. A dynamic neural model of tonal downstep. Paper presented to the Proceedings of the Annual Meetings on Phonology, 2025a.

Chaturvedi, Manasvi & Jason A. Shaw. 2025b. Phonetic trace effects in experimentally-induced vowel errors. Cognitive Neuropsychology 42.106-26.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. Journal of Phonetics 54.35-50.

Davy, Jim & Derek Nurse. 1982. Synchronic versions of Dahl's Law: The multiple applications of a phonological dissimilation rule. Journal of African Languages and Linguistics 4.157-95.

Elie, Benjamin, David N Lee & Alice Turk. 2023. Modeling trajectories of human speech articulators using general Tau theory. Speech Communication 151.24-38.
Ellis, Lucy & William J Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. Journal of Phonetics 30.373-96.
Erlhagen, Wolfram & Gregor Schöner. 2002. Dynamic field theory of movement preparation. Psychological Review 109.545-72.
Fromkin, Victoria. 1971. The non-anomalous nature of anomalous utterances. Language 47.27-52.
Gafos, Adamantios I & Christo Kirov. 2010. A dynamical model of change in phonological representations: The case of lenition. Approaches to phonological complexity, ed. by F. Pellegrino, E. Marsico, I. Chitoran & C. Coupé, 225-46. Berlin: Mouton de Gruyter.
Goldsmith, John & Gary Larson. 1990. Local modeling and syllabification. Papers from the 26th Regional Meeting of the Chicago Linguistic Society, Volume 2: The Parasession on the Syllable in Phonetics and Phonology, ed. by M. Ziolkowski, M. Noske & K. Deaton, 129-42. Chicago: Chicago Linguistic Society.
Goldstein, Louis. 2011. Back to the past tense in English. Representing language: Essays in honor of Judith Aissen.69-88.
Gurevich, Naomi. 2013. Lenition and contrast: The functional consequences of certain phonetically conditioned sound changes: Routledge.
Hargus, Sharon. 2018. The lexical phonology of Sekani: Routledge.
Iskarous, Khalil. 2017. The relation between the continuous and the discrete: A note on the first principles of speech dynamics. Journal of Phonetics 64.8-20.
Iskarous, Khalil & Louis Goldstein. 2018. The dynamics of prominence profiles: from local computations to global patterns. Shaping phonology.253-77.
Jordan, Michael Irwin. 1985. The learning of representations for sequential performance (connectionism, parallel models, serial order): University of California, San Diego.
Katsika, Argyro, Jelena Krivokapić, Christine Mooshammer, Mark Tiede & Louis Goldstein. 2014. The coordination of boundary tones and its interaction with prominence. Journal of Phonetics 44.62-82.
Kelso, JA Scott, Elliot L Saltzman & Betty Tuller. 1986. The dynamical perspective on speech production: Data and theory. Journal of Phonetics 14.29-59.
Kim, Seung-Eun & Sam Tilsen. 2025. The Gesture-Field-Register (GFR) framework for modeling F0 control. Journal of Phonetics 110.101410.
Kirkham, Sam. 2025. Discovering dynamical laws for speech gestures. Cognitive Science 49.e70064.
Kirkham, Sam & Patrycja Strycharczuk. 2024a. A dynamic neural field model of vowel diphthongisation. Paper presented to the Proceedings of the 13th International Seminar on Speech Production, 2024a.
—. 2024b. A dynamic neural field model of vowel diphthongisation. Proc. ISSP.193-96.
Kirov, Christo & Adamantios Gafos. 2007. Assembling phonological representations. Phonological Systems and Complex Adaptive Systems: Phonology and Complexity). Berlin/New York: Mouton de Gruyter.
Kochetov, Alexei & Marianne Pouplier. 2008. Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. Phonology 25.399-431.
Mielke, Jeff. 2005. Ambivalence and ambiguity in laterals and nasals. Phonology 22.169-203.
Mizoguchi, Ai, Ayako Hashimoto, Sanae Matsui, Setsuko Imatomi, Ryunosuke Kobayashi & Mafuyu Kitahara. 2020. Neutralization of Voicing Distinction of Stops in Tohoku Dialects of Japanese: Field Work and Acoustic Measurements. Proceedings of Interspeech.1873-77.
Nadeu, Marianna & José Ignacio Hualde. 2015. Biomechanically conditioned variation at the origin of diachronic intervocalic voicing. Language and Speech 58.351-70.
Nam, Hosung, Louis Goldstein & Elliot Saltzman. 2009. Self-organization of syllable structure: a coupled oscillator mode. Approaches to phonological complexity, ed. by F. Pellegrino, E. Marisco & I. Chitoran, 299-328. Berlin, New York: Mouton de Gruyter.
Nowak, Martin A, Natalia L Komarova & Partha Niyogi. 2002. Computational and evolutionary aspects of language. Nature 417.611-17.
Parrell, Benjamin & Shrikanth Narayanan. 2018. Explaining coronal reduction: Prosodic structure and articulatory posture. Phonetica 75.151-81.
Pierrehumbert, Janet, Mary E Beckman & D Robert Ladd. 2000. Conceptual foundations of phonology as a laboratory science. Phonological knowledge: Conceptual and empirical issues.273-304.
Pintado-Urbanc, Alessandra. 2025. A dynamic neural field model of asymmetric interference effects in code-switching. Proceedings of the Linguistic Society of America 10.5909-09.
Roon, Kevin D & Adamantios I Gafos. 2016. Perceiving while producing: Modeling the dynamics of phonological planning. Journal of Memory and Language 89.222-43.
Saltzman, E. & K.G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. Ecological Psychology 1.333-82.
Sande, Hannah & Madeleine Oakley. 2023. A typological survey of the phonological behavior of implosives: Implications for feature theories. Phonological Data and Analysis 5.1–46-1–46.

Schneegans, Sebastian. 2021. COSIVINA: A Matlab Toolbox to Compose, Simulate, and Visualize Neurodynamic Architectures (Version 1.4).
Schöner, Gregor & John P Spencer. 2016. Dynamic thinking: A primer on dynamic field theory Oxford: Oxford University Press.
Shaw, Jason A. 2025. Unifying phonological and phonetic aspects of speech in dynamic neural fields: The case of laryngeal patterns in Japanese. Phonological Studies 28.1-12.
Shaw, Jason A & Shigeto Kawahara. 2023. Limits on gestural reorganization following vowel deletion: The case of Tokyo Japanese. Laboratory Phonology 14.
Shaw, Jason A & Kevin Tang. 2023. A dynamic neural field model of leaky prosody: proof of concept. Paper presented to the Proceedings of the Annual Meetings on Phonology, 2023.
Smolensky, Paul & Matthew Goldrick. 2016. Gradient symbolic representations in grammar: The case of French liaison. Rutgers Optimality Archive 1552.1-37.
Sonderegger, Morgan & Partha Niyogi. 2013. Variation and Change in English Noun/Verb Pair Stress: Data. Dynamical Systems Models, and Their Interaction. In ACL Yu, ed., Origins of Sound Patterns: Approaches to Phonologization.262-84.
Sorensen, Tanner & Adamantios Gafos. 2016. The gesture as an autonomous nonlinear dynamical system. Ecological Psychology 28.188-215.
Stern, Michael. 2025. Dynamic Field Theory unifies discrete and continuous aspects of linguistic representations. Proceedings of the Linguistic Society of America 10.5934-34.
Stern, Michael C, Manasvi Chaturvedi & Jason A Shaw. 2022. A dynamic neural field model of phonetic trace effects in speech errors. Paper presented to the Proceedings of the Annual Meeting of the Cognitive Science Society, 2022.
Stern, Michael C & Jason A Shaw. 2023. Neural inhibition during speech planning contributes to contrastive hyperarticulation. Journal of Memory and Language 132.104443.
—. 2025. Nonlinear second-order dynamics describe labial constriction trajectories across languages and contexts. Journal of Phonetics 111.101427.
Stevens, Kenneth. 1989. On the quantal nature of speech. Journal of Phonetics 17.3-45.
Strogatz, Steven H. 2024. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering: Chapman and Hall/CRC.
Suzuki, Keiichiro. 1998. A Typological Investigation of Dissimilation. Tucson, AZ: University of Arizona Ph. D. Dissertation.
Teng, Xiangbin, Xing Tian, Jess Rowland & David Poeppel. 2017. Concurrent temporal channels for auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. PLoS biology 15.e2000812.
Tiede, Mark. 2005. MVIEW: software for visualization and analysis of concurrently recorded movement data. New Haven, CT: Haskins Laboratories.
Tilsen, Sam. 2019. Motoric mechanisms for the emergence of non-local phonological patterns. Frontiers in psychology 10.2143.