Categorical alternations from neural field dynamics

Background: Dynamic Neural Fields (DNF)'s have been effectively deployed to model the time course of speech planning, including subtle shifts in phonetic targets under the influence of multiple phonological inputs into the speech planning processes (Stern & Shaw 2023ab, Shaw & Tang 2023; Stern, Chaturvedi, Shaw 2022; Roon & Gafos 2016). However, most of these models have focused only on the planning of very simple phonological forms, i.e., single segments or single features. Here, we expand these models to plan phonological sequences, represented as temporally-spaced inputs to a DNF. We show that changes in DNF activation in response to one input can impact how the field evolves in response to subsequent inputs. Specifically, an activation peak in one location of the DNF inhibits activation in distal locations of the field. Inhibition of this sort can prevent subsequent inputs from driving the field across activation threshold for production. Variation in the timing and amplitude of inputs to the DNF derives a typology of patterns typically characterized as phonological alternations.

Case study: we illustrate this property of DNFs with laryngeal phonology. For an input of /$C_{[+sg]}VC_{[+sg]}V$/, where V indicates a vowel (voiced by default) and $C_{[+sg]}$ indicates a consonant specified for a spread glottis feature, we illustrate how the timing and amplitude of inputs to a glottal width DNF can derive three different patterns: (1) vowel devoicing [$C_{[+sg]}$ V̥$C_{[+sg]}$V]; (2) intervocalic voicing [$C_{[+sg]}$ $VC_{[+voi]}$V]; and (3) and a fully faithful voicing contour [$C_{[+sg]}VC_{[+sg]}$V]. All three of these patterns are attested cross-linguistically. For concreteness, we parameterized our models with consideration of dialect variation within Japanese: Tokyo Japanese shows high vowel devoicing, pattern (1); Osaka Japanese shows faithful realization of laryngeal features, pattern (2); and Tohoku Japanese shows intervocalic voicing, pattern (3) (Fujimoto et al. 2002; Mizuguchi et al., 2020).

Simulations: We derived the three patterns from the timing and amplitude of inputs to a glottal width DNF. The DNF was 100 units wide, representing neuron populations controlling degrees of glottal width, ranging from 0 (closed) to 100 (maximally spread glottis). Change in activation across the field over time is dictated by the differential equation below, where u is activation, h is resting activation (h = -5 for all simulations), s is input, k is the interaction kernal, and q is noise. Inputs take the form of gaussian distributions with three parameters: field position, p; width, w; and amplitude, a. The spread glottis input, $s_{[sg]}(x,t)$, was centered on 75 (p = 75); the voicing input, $s_{[voicing]}(x,t)$, was centered on 25 (p = 25). For each input, w = 20. The amplitude, a, for the spread glottis input was held constant at 7 (c.f., resting level of -5); the amplitude for the voicing input was held constant at 5 (just enough to counter resting level of -5). These parameters were selected to model the assumption that voicing is the default state of the glottis (Browman & Goldstein 1989) by having a constant voicing input amplitude of 5 across all segments. All else equal, the strong input, a = 7, for spread glottis overpowers the default voicing state, leading to a peak at the wide end of the field. The interaction kernal, k, was parameterized for local excitation and global inhibition. Gated by a sigmoid, g, the interaction kernal impacts the field only when activation in some field location crosses threshold (u = 0), functioning to stabilize a local activation peak to serve as speech production target. The timing of inputs to the field was dictated by a mora-based clock. If the duration of the [sg] inputs were relatively long, i.e., half a mora cycle (Figure 1), a peak at the voicing end of the field would fail to cross threshold before the second [sg] input entered the field, leading to vowel devoicing. Shortening the duration of the [sg] input to ¼ of the mora cycle and suppressing default voicing during the second [sg] input leads to faithful realization of a voicing contour (Figure 2). If the default voicing input is not suppressed, then the second [sg] input will fail to cross threshold, leading to intervocalic voicing (Figure 3). These three patterns derive from

inputs that can be all characterized as /C[+sg]VC[+sg]V/ yet they differ in the relative timing and duration of the inputs to speech production planning, which has categorical implications for the surface form.

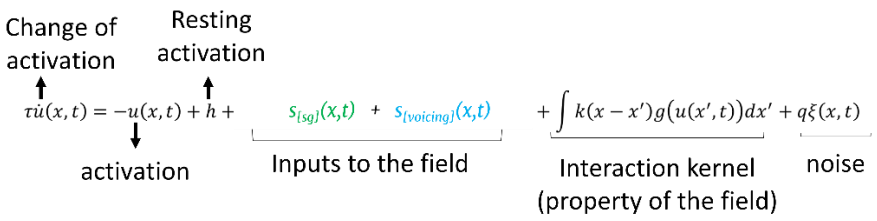
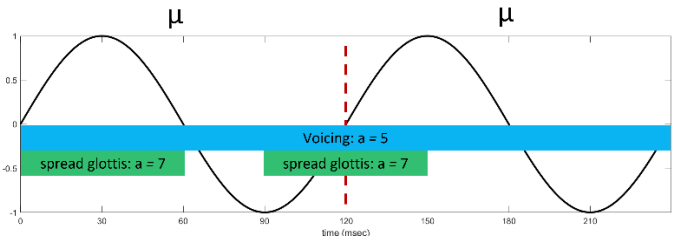


Figure 1. Vowel devoicing (right) occurs when [sg] inputs (below) are ½ of the mora cycle and the second mora [sg] input overlaps the first. Blue indicates voicing; green indicates voicelessness.



/CVCV/→[CV̥CV]

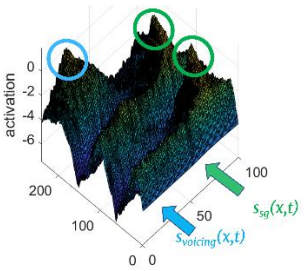
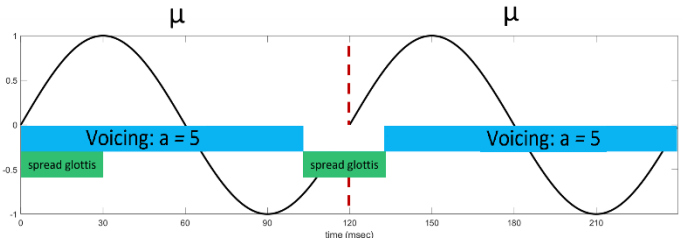


Figure 2. faithful voicing contour (right) occurs when [sg] inputs (below) are ¼ of the mora cycle, the second mora [sg] input overlaps the first, and default voicing is suppressed during the second spread glottis input.



/CVCV/→ [CVCV]

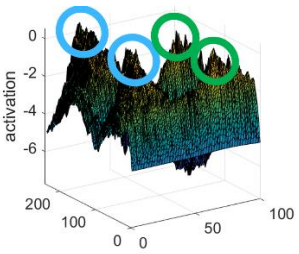
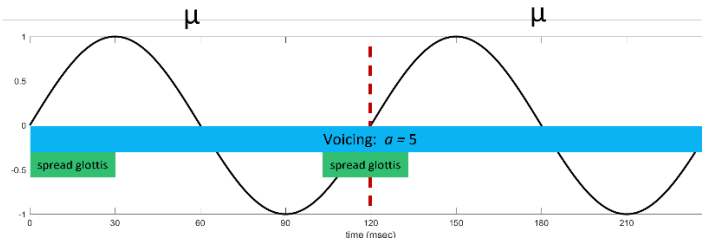
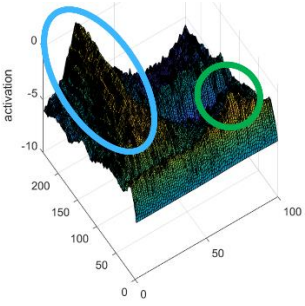


Figure 3. intervocalic voicing (right) occurs when [sg] inputs (below) are ¼ of the mora cycle, the second mora [sg] input overlaps the first, and default voicing is not suppressed.



/CVCV/→ [CVCV]



References: Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201-251. **Fujimoto, M., Murano, E., Niimi, S., & Kiritani, S. (2002).** Differences in glottal opening pattern between Tokyo and Osaka dialect speakers: factors contributing to vowel devoicing. *Folia phoniatrica et logopaedica*, *54*(3), 133-143. **Mizoguchi, A., Hashimoto, A., Matsui, S., Imatomi, S., Kobayashi, R., & Kitahara, M. (2020).** Neutralization of Voicing Distinction of Stops in Tohoku Dialects of Japanese: Field Work and Acoustic Measurements. In *INTERSPEECH* (pp. 1873-1877). **Roon, K. D., & Gafos, A. I. (2016).** Perceiving while producing: Modeling the dynamics of phonological planning. *Journal of Memory and Language*, *89*, 222-243. **Shaw, J. A., & Tang, K. (2023)**. A dynamic neural field model of leaky prosody: proof of concept. In *Proceedings of the Annual Meetings on Phonology* (Vol. 10). **Stern, M. C., Chaturvedi, M., & Shaw, J. A. (2022).** A dynamic neural field model of phonetic trace effects in speech errors. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44). **Stern, M. C., & Shaw, J. A. (2023)**. Neural inhibition during speech planning contributes to contrastive hyperarticulation. *Journal of Memory and Language*, *132*, 104443.