Parameter dynamics for articulatory gestures

Jason A. Shaw[1], Sejin Oh[2]

[1]Yale University (USA); [2]Hanyang Institute for Phonetics and Cognitive Science (Korea)

Synopsis: Articulatory Phonology makes use of dynamically-defined gestures to model how phonological targets drive movement of articulators over time. Invariance is found at the level of the gesture parameters, *stiffness* and *target,* which impart invariance to the gesture. Since multiple gestures can overlap in time and exert joint influence on articulators, invariance at the level of the gesture can still predict context-specific kinematic variation through gestural blending (e.g., Saltzman & Munhall 1989; Browman & Goldstein 1990; Marin 2007; Iskarous et al. 2012). Gestural blending, however, cannot account for all context-specific kinematic variation. Motivated by one such case, that of assimilatory palatalization in Russian (Oh 2022; Oh et al. 2024), we propose a parameter dynamics for articulatory gestures, whereby the parameters of the gestures are not invariant but vary lawfully according to a dynamical system. We formalize the proposal using Dynamic Field Theory (e.g., Schöner & Spencer 2016), following recent work (Roon & Gafos 2016; Harper 2021; Stern et al 2022; Stern & Shaw 2023ab; Shaw & Tang 2023).

Phenomenon: Russian contrasts plain and palatalized consonants with the “plain” consonants having velarization/uvularization, i.e., /pˠ/ vs. /pʲ/ (e.g., Avanesov 1972, Timberlake 2004, Roon & Whalen 2019). In stop-glide sequences, e.g., /pˠj/, the contrast is neutralized, due to palatalization of the plain stop, e.g., /pˠjot/→[pʲjot] (assimilatory palatalization). Using Electromagnetic Articulography (EMA), Oh et al. (2024) showed that contrast neutralization is incomplete. Underlyingly plain stops are palatalized, but they show small but significant differences from underlyingly palatalized segments. Specifically, the tongue body (TB) position at the onset of movement towards the palate was more retracted and the TB movement started later in time (Figure 3). Oh et al. (2024) concludes that incomplete neutralization follows from some residual presence of an underlying velar TBCL (tongue blade constriction location) gesture for plain stops.

Static parameters: Varying the blending strength of two overlapping TBCL gestures, one with a velar target and one with a palatal target, produces a range of kinematic outcomes, but this approach cannot produce the observed kinematics of assimilatory palatalization. As shown by Oh (2022), deriving the spatial difference found at the onset of TB advancement requires that a velar gesture first drives the TB back (retraction) *before* the TB advances. This is impossible if the gestures are in-phase and have fixed gestural parameters, including target, blending strength, and stiffness. We instead derived the result by implementing a dynamics for gesture parameters. Instead of blending two TBCL gestures, we have one TBCL gesture whose target varies over time.

Parameter dynamics: We propose a dynamic neural field (DNF) representing TB constriction location (TBCL) which evolves under the influence of a velar input and a palatal input. Local excitation and global inhibition within the DNF ensures that despite multiple inputs, only one activation peak (corresponding to a CL target) forms at any time. For assimilatory palatalization (Figure 2), the field stabilizes first on the velar end of the TBCL field; this triggers a polarity reversal of the inputs from excitation/inhibition to inhibition/excitation (e.g.. Stern & Shaw 2023b), which allows the velar peak to dissipate and a new peak to form at the palatal end of the field. Figures 1 & 2 show single simulation runs (DNF only); Figure 5 shows the gestural dynamics, where the target of the gesture is dictated by the location of activation peaks in the DNF; across 100 runs, the dynamic target–evolving over time from velar to palatal–reliably produces the differences (incomplete neutralization) observed in the data (Figure 4). More broadly, on this approach, invariance is found not in the parameters themselves but in the dynamics of the DNFs that condition change in gesture parameter activation over time.

References: Avanesov, R. I. (1972). Russkoe literaturnoe proiznosenie. [Literary pronunciation of Russian.] Moscow: Prosvescenie. **Browman, C. P., & Goldstein, L. (1990).** Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics, 18*(3), 299-320. **Harper, S. K. (2021).** *Individual differences in phonetic variability and phonological representation* (Doctoral dissertation, University of Southern California). **Iskarous, K., McDonough, J., & Whalen, D. H. (2012).** A gestural account of the velar fricative in Navajo. *Laboratory Phonology, 3*(1), 195-210. **Marin, S. N. (2007)**. *Vowel to vowel coordination, diphthongs and Articulatory Phonology*. Yale University. **Oh, S., Shaw, J. A., Durvasula, K., & Kochotov, A. (2024).** Russian assimilatory palatalization is incomplete neutralization. *Laboratory Phonology, arXiv preprint arXiv:2306.12789*.

Oh, S. (2022). *Incomplete neutralization in articulatory phonology* (Doctoral dissertation, City University of New York). **Roon, K. D., & Gafos, A. I. (2016)**. Perceiving while producing: Modeling the dynamics of phonological planning. *Journal of Memory and Language*, *89*, 222-243. **Roon, K. D., & Whalen, D. H. (2019)**. Velarization of Russian labial consonants. In *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 3488-3492). **Saltzman, E. L., & Munhall, K. G. (1989)**. A dynamical approach to gestural patterning in speech production. *Ecological psychology*, *1*(4), 333-382. **Schöner, G., & Spencer, J. P. (2016)**. *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press. **Shaw, J. A., Oh, S., Durvasula, K., & Kochetov, A. (2021)**. Articulatory coordination distinguishes complex segments from segment sequences. *Phonology*, *38*(3), 437-477. **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. (2023a)**. Neural inhibition during speech planning contributes to contrastive hyperarticulation. *Journal of Memory and Language*, *132*, 104443. **Stern, M.C. & Shaw, J. A. (2023b)**. Not all phonological neighbors affect production equivalently: predictions from a neural dynamic model. *Proceedings of the 20th International Congress of Phonetic Sciences*. Prague, Czech Republic. **Timberlake, A. (2004)**. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Figure 1. DNF activation over time (one simulation run) for the underlying palatal. Peak at the palatal (front) end of the field remains above threshold throughout the gesture.

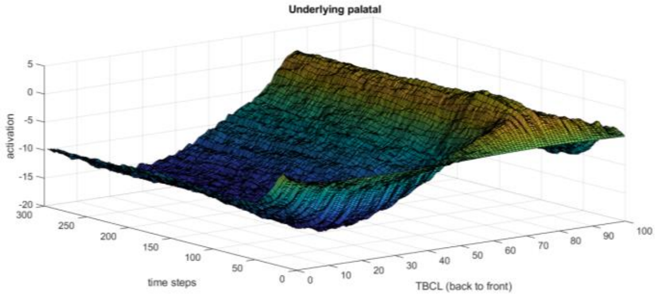

Figure 2. DNF activation over time for an assimilatory palatal. Peak at the velar (back) end of the field emerges and then dissipates and a new peak forms at the palatal end (front), changing TBCL target from velar to palatal.

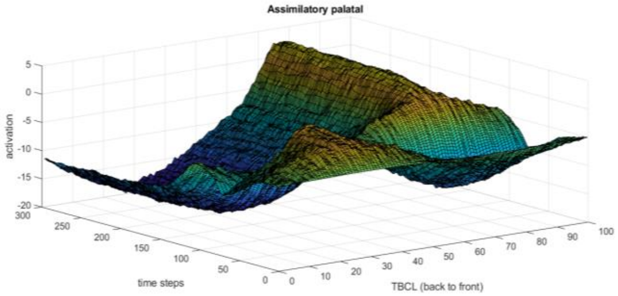

Figure 3. Data from Oh et al. (2024) comparing TB position at gesture onset (left) and onset lag, i.e., when the palatal movement starts relative to the lip movement (right), across conditions.

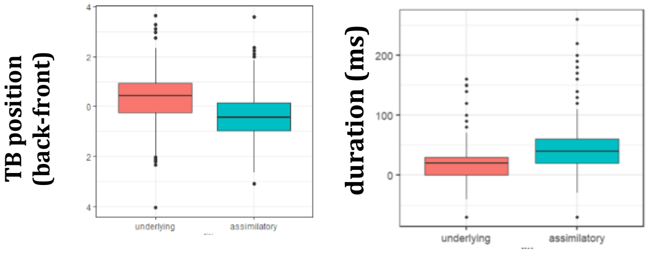

Figure 4. Results from 100 simulation runs comparing TB position at gesture onset (left) and onset lag (right) across conditions, c.f., corresponding data in Fig 3.

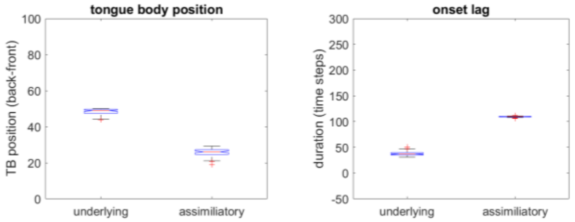

Figure 5 (right→). Kinematic trajectories of the tongue body simulated from three different TBCL targets: velar (for the "plain" consonants), palatal, and assimilatory palatal.