Linguistic Research 43(2): 497-534
DOI: 10.17250/khisli.43.2.202606.007

Articulatory modeling of incomplete neutralization in Russian palatalization*

Sejin Oh
(Jeju National University)

Oh, Sejin. 2026. Articulatory modeling of incomplete neutralization in Russian palatalization. *Linguistic Research* 43(2): 497-534. Russian contrasts plain and palatalized consonants, but this contrast is commonly reported to be neutralized before a palatal glide, where plain consonants undergo assimilatory palatalization. Recent articulatory evidence, however, suggests that underlying and assimilatory palatalization may be phonologically neutralized, but it is phonetically incomplete. This study investigates the internal gestural organization that gives rise to such incomplete neutralization in Russian. Building on the empirical findings of previous work, the current study employs articulatory modeling to examine how different gestural coordination can generate subtle differences between the two palatalization types. We created four gestural models using Task Dynamic Application (TADA). Each model varied in the presence or absence of velar gestures and in the coordination relations among labial, palatal, and velar articulatory gestures. Simulation outputs were directly compared with temporal and spatial measures reported in previous work. The model that best accounted for the empirical data posits eccentric-phase coordination between velar and palatal gestures in assimilatory palatalization, combined with in-phase coupling between labial and palatal gestures. This model successfully replicated the empirical findings: complex-segment timing for both conditions as well as greater tongue body retraction at palatal gesture onset and longer onset-to-onset lags in assimilatory palatalization. The current study explicitly shows the internal gestural organization that can account for incomplete neutralization between underlying and assimilatory palatalization in Russian. **(Jeju National University)**

Keywords computational modeling, Task Dynamic Application (TADA), Russian palatalization, incomplete neutralization, gestural coordination

* This work was supported by the research grant of Jeju National University in 2026.

1. Introduction

Russian contrasts palatalized and plain (non-palatalized) consonants, but this contrast is reported to be neutralized when a plain consonant is followed by a glide (e.g., Kochetov 2011). That is, both a palatalized consonant and a plain consonant preceding a palatal glide are realized as a palatalized consonant. For example, a plain consonant preceding a palatal glide in /pjot/ is realized as a palatalized consonant [pʲ], resulting in neutralization of the contrast between plain and palatalized consonants in Russian, as shown in (1). In this study, palatalized consonants are treated as underlying palatalization, whereas plain consonants preceding a palatal glide are analyzed as instances of assimilatory palatalization. However, previous studies have reported that these "plain" consonants possibly have a secondary articulation involving retraction of the tongue dorsum (velarization/uvularization, see Litvin 2014; Roon and Whalen 2019). For example, a recent ultrasound study by Roon and Whalen (2019) confirmed that plain consonants in Russian are velarized (and/or uvularized) with intra-speaker variation. Building on these findings, recent work suggests that the two types of palatalization may not be fully identical in their articulatory realization (Oh et al. 2024).

(1)	Palatalized segments (Underlying palatalization)	Consonants preceding a palatal glide (Assimilatory palatalization)
	/pʲatij/ [pʲatɨj] 'fifth'	vs. /pjanij/ [pʲjanɨj] 'drunk'
	/pʲok/ [pʲok] 'bake (3ps past)'	vs. /pjot/ [pʲjot] 'drink (3ps pres)'
	/bjust/ [bjust] 'bust'	vs. /bjut/ [bjjut] 'beat (3p pl)'

In Oh et al. (2024), it was hypothesized that the gestural blending of two secondary articulation gestures (palatalization and velarization/ uvularization) in the assimilatory palatalization condition leads to incomplete neutralization of underlying and assimilatory palatalization in Russian. Experimental evidence from EMA supports this hypothesis in that underlying and assimilatory palatalization exhibit inter-gestural coordination characteristic of complex segments, while residual evidence of an underlying tongue dorsum retraction (velarization/uvularization) gesture remains in the assimilatory palatalization condition. More specifically, a key finding from Oh et al. (2024) is that both underlying and assimilatory palatalizations show no

correlation between consonant duration and onset-to-onset lag, a pattern consistent with complex segmental coordination. This result suggests that the contrast between palatalized and plain consonants is neutralized to the palatal counterpart when a plain consonant is followed by a glide. Crucially, however, this neutralization is phonetically incomplete. Oh et al. (2024) found that the tongue body was significantly more retracted for assimilatory palatalization than for underlying palatalization at the onset of the palatal gesture, although the difference was small (1.5 mm). In addition, onset-to-onset lag was significantly longer for assimilatory palatalization than for underlying palatalization. Together, these small but significant differences indicate incomplete neutralization of the palatal–plain contrast and align with previous findings that Russian plain consonants exhibit secondary velarization.

Oh et al. (2024) also discuss an alternative account in which the tongue body differences across palatalization types arise from differences in vowel–consonant coordination rather than from incomplete neutralization. On this view, assimilatory palatalization involves earlier initiation of the back vowel gesture, leading to greater tongue retraction at the onset of the palatal gesture via gestural blending. However, this account predicts segment-sequence–like timing, which is inconsistent with the EMA evidence for complex-segment coordination (see Section 5.3 in Oh et al. 2024 for details). With respect to the delayed onset of the palatal tongue body movement in assimilatory palatalization, Oh et al. (2024) consider several possibilities within an articulatory phonology framework. They argue that differences in blending strength alone, as well as competitive (anti-phase) coupling, fail to account for the observed temporal patterns. Instead, they suggest that an eccentrically timed velar gesture, initiated slightly earlier than the palatal gesture while preserving complex-segment coordination, could jointly account for both the spatial and temporal asymmetries observed in the EMA data (see Section 5.4 in Oh et al. 2024 for details).

While these findings and considerations favor an analysis involving complex segment coordination with residual effects of an underlying dorsal retraction gesture, the precise internal gestural organization responsible for these patterns remains underspecified. This gap motivates the present study. Using gestural modeling, the current study systematically explores how different internal gestural configurations can give rise to the temporal coordination patterns and subtle articulatory differences reported in Oh et al. (2024). By evaluating multiple modeling scenarios within a coupled oscillator framework, the study seeks to identify gestural representations that

are compatible with both the EMA-based temporal evidence for complex segments and the articulatory asymmetries between underlying and assimilatory palatalization. To this end, several gestural models are implemented in Task Dynamic Application (TADA), and their simulation outputs are directly compared with the EMA findings.

The remainder of the paper is organized as follows. Section 2 introduces the modeling framework: Section 2.1 outlines Articulatory Phonology, Section 2.2 presents the Task Dynamic Application (TADA), and Section 2.3 reviews methods for incorporating temporal variation in gestural simulations. Section 3 presents four gestural models designed to reproduce the EMA results and details their gestural specifications and coupling relations. Section 4 reports the results of the TADA simulations, and Section 5 provides discussion and concluding remarks.

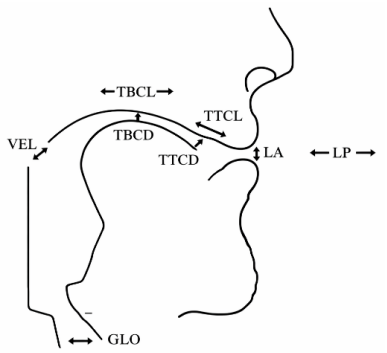
2. Modeling framework

2.1 Articulatory Phonology

In the Articulatory Phonology (henceforth, AP) framework, gestures, which are abstract representations of movement of articulators in the vocal tract, serve as the primitive phonological units (e.g., Browman and Goldstein 1986, 1989, 1992, 1995; Pouplier 2020). Speech can be organized into constellations of gestures, and each gesture is defined as an event which forms and releases a constriction in the vocal tract. Crucially, gestures are specified spatially as well as temporally. AP framework modeled by a dynamical system is different from classical phonological representations, in that it bridges abstract phonological representations and continuous physical movement.

Gestures are discrete and abstract in the sense that they are specifically defined by a set of dynamical parameters which characterize each gesture distinctively (e.g., Browman and Goldstein 1986, 1989, 1992, 1995; Pouplier 2020). As shown in Figure 1, gestures are specified with respect to vocal tract variables. AP utilizes a set of gestural descriptors which distinguish contrastive gestures: Constriction degree (CD), constriction location (CL), and stiffness (k). Tract variable goals (input values for CD and CL) determine the inherent spatial aspect, while the stiffness specifies the intrinsic temporal aspect of each gesture. Values for the possible descriptor are shown

in (2) (adapted from Browman and Goldstein 1989: 207). For example, /s/ and /ʃ/ differ in their values for CL (alveolar vs. postalveolar, respectively), and /s/ and /t/ differ in their values for CD (critical vs. closed).



Tract variable		Articulators involved
LP	Lip protrusion	Upper & lower lips, jaw
LA	Lip aperture	Upper & lower lips, jaw
TTCL	Tongue tip constrict location	Tongue tip, body, jaw
TTCD	Tongue tip constrict degree	Tongue tip, body, jaw
TBCL	Tongue body constrict location	Tongue body, jaw
TBCD	Tongue body constrict degree	Tongue body, jaw
VEL	Velic aperture	Velum
GLO	Glottal aperture	Glottis

Figure 1. Tract variables (adapted from Browman and Goldstein 1989: 207)

(2) CD descriptors: closed, critical, narrow, mid, wide
CL descriptors: protruded, labial, dental, alveolar, postalveolar, palatal, velar, uvular, pharyngeal

In the AP framework, gestures are defined not only by their spatial targets but also by their temporal coordination, and this temporal dimension becomes especially crucial when multiple gestures combine within complex syllable structures. Building on this perspective, the coupled oscillator model proposed distinct coordination patterns depending on syllable structure (Goldstein et al. 2006; Saltzman et al. 2008; Goldstein et al. 2009; Nam et al. 2009). For example, in CV syllables, consonantal and vocalic gestures are timed in-phase, meaning that the gestures start simultaneously (relative phase of 0°). By contrast, in VC and CC syllables, gestures are coordinated in anti-phase, reflecting a sequential timing relationship in which the initiation of one gesture precedes the other (relative phase of 180°). When it comes to CCV syllables, the coordination of gestures within complex onsets and the vocalic gesture

involves a competitive coupling: each consonant is coupled in-phase with the vowel, while simultaneously maintaining anti-phase relations with one another. This interplay yields the so-called c-center effect, whereby the vowel is initiated at the temporal midpoint of the consonant cluster. The c-center effect has been documented in languages like English (Browman and Goldstein 1988), French (Kühnert et al. 2006), Georgian (Goldstein et al. 2007), Serbian (Tilsen et al. 2012), etc. In contrast, other languages exhibit simple onset timing, where consonants are coupled anti-phase with each other but only the right-most consonant is in-phase with the vowel. This pattern results in a sequential alignment rather than a centralized onset, and has been observed in Moroccan Arabic (Shaw et al. 2011), Tashlhiyt Berber (Hermes et al. 2017), Hebrew (Tilsen et al. 2012), Montreal French (Tilsen et al. 2012), etc.

2.2 Task Dynamic Application (TADA)

Task Dynamic Application (henceforth, TADA) is a MATLAB-based software for simulating the gestural representations of utterances and generating acoustic output (Nam et al. 2004, 2006, 2012). Within the Articulatory Phonology framework, TADA operationalizes the theory by translating abstract gestural specifications into dynamic trajectories of articulatory movement, thereby linking phonological representations with measurable acoustic signals. Importantly, TADA is not merely a speech synthesizer but a computational instantiation of hypotheses about speech planning and timing. Based on the Task Dynamic model of inter-articulator coordination in speech (e.g., Saltzman and Munhall 1989), TADA implements the following models that feed one another as shown in Figure 2.

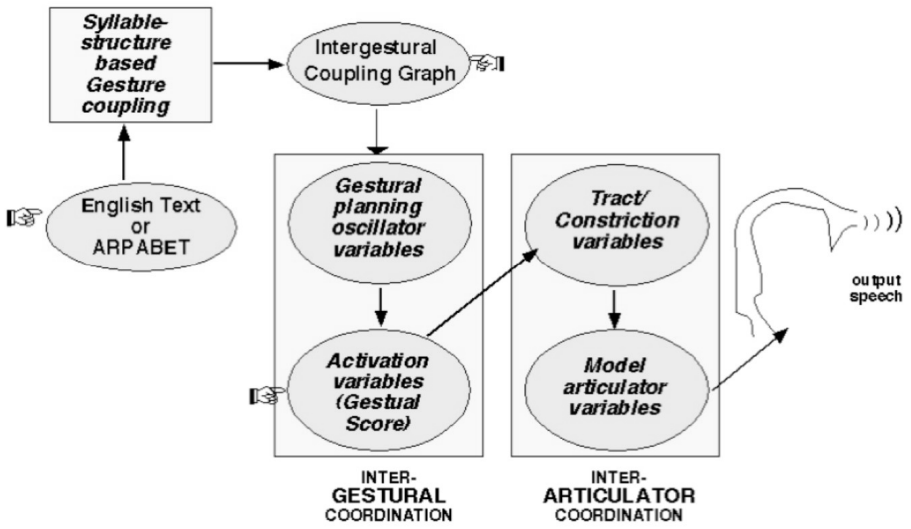


Figure 2. Information flow through TADA models (adapted from Nam et al. 2006: 2)

First, with a text string as input, the Syllable structure-based gesture coupling model generates an intergestural coupling graph, which feeds into the coupled oscillator model of inter-gestural coordination. In turn, the coupled oscillator model generates a gestural score, which specifies inter-gestural coordination and becomes an input of the task dynamic model of inter-articulator coordination. Then, the task dynamic model generates the vocal tract constriction variables and the articulatory degrees of freedom. The outcome feeds into Configurable Articulatory Synthesis (CASY) to compute a time-varying vocal tract area function and the resonance frequencies and bandwidths corresponding to those area functions. Finally, taking these as an input, Sensimetrics' HLsyn synthesizer generates the acoustic output. Notably, these models can run separately and independently when an input is provided for each model.

In TADA, coupling graphs can be created through the GEST menu by using either English text or ARPABET as an input, resulting in TV<id>.O and PH<id>.O files. TV<id>.O file contains gestural specifications. It consists of the specification of control parameters of each gesture and a label specifying the oscillator that controls the activation of that gesture (corresponds to an oscillator label in the PH.O file). PH<id>.O file, on the other hand, contains timing oscillator and coupling

specifications. It consists of dynamical parameter values, the phases for activation and deactivation of gestures, and the coupling parameters for oscillator pairs.

As an example, TVbutte.O and PHbutte.O files are created through the GEST menu in TADA for the English word 'butte.' As shown in Figure 3, TVbutte.O consists of a list of gestures and their positions within syllables, and each line contains dynamical parameters, articulator weights, and blending parameters. As shown in lines 18-21 in Figure 3, the second onset /j/ in the word butte contains three tract variables. Two are for the TB gesture, one specifying the constriction degree (TBCD) the other the constriction location (TBCL). There is also a gesture for lip aperture (LA).

Each line in Figure 3 consists of the following information: TV_name Osc_ID target freq damp art_wts alpha beta. In line 21, for example, TV_name is 'TBCL,' and Osc_ID is 'ons2_nar1' which corresponds to an oscillator label in the PHbutte.O file (see line 6 in Figure 4). target is 95 degrees since the target specification for palatal is 95 degrees (c.f., target specifications for CD is in mm; e.g., see line 20). freq specifies the stiffness of a gesture in which the default is 8 Hz for gestures associated with consonants (c.f., the default for vowels is set to 4 Hz; e.g., see lines 25-28). This stiffness parameter is used as one of the ways to elicit temporal variation in Section 4.2.2. damp refers to the Damping ratio which is set to 1 by default. art_wts specifies an articulator weight where the higher value indicates that the articulator is "heavy" and less likely to move in the production of a constriction. alpha specifies the blending strength of the gesture. A higher value indicates a stronger blending strength. In line 21, alpha is set to 100 beta is set to the reciprocal of alpha, 0.01 (c.f., when alpha is zero, beta is also set to zero; e.g., see line 15).

```
10 0 7


% Input string: <butte>
%
%
% Word 1:   butte
% arpabet:  (B Y-UW1_W T)
%
%
% syllable 1:   B Y-UW1_W T
%
%    onset cluster = <B Y>
%    segment 1 [B]:
'VEL' 'ons1_clo1' -0.1 8 1 NA=1 0 0
'LA' 'ons1_clo1' -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 'ons1_rel1' 11 8 1 JA=8,UH=5,LH=1 1 1
%    segment 2 [Y]:
'LA' 'ons2_nar1' 8 8 1 JA=8,UH=5,LH=1 1 1
'TBCD' 'ons2_nar1' 2 8 1 JA=10,CL=1,CA=1 100 0.01
'TBCL' 'ons2_nar1' 95 8 1 JA=10,CL=1,CA=1 100 0.01
%
%    nucleus cluster = <UW1>
%    segment 1 [UW]:
'LP' 'v_rnd1' 12 4 1 LX=1 1 1
'TBCL' 'v1' 125 4 1 JA=10,CL=1,CA=1 1 1
'TBCD' 'v1' 4 4 1 JA=1,CL=1,CA=1 1 1
'LA' 'v_rnd1' 5 4 1 JA=1,UH=5,LH=1 1 1
%
%    coda cluster = <W T>
%    segment 1 [W]:
'LA' 'cod1_rel1' 11 8 1 JA=8,UH=5,LH=1 1 1
'LA' 'cod1_nar1' 1 8 1 JA=8,UH=5,LH=1 1 1
%    segment 2 [T]:
'TTCL' 'cod2_rel1' 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'GLO' 'cod2_h1' 0.4 16 1 GW=1 0 0
'TTCD' 'cod2_clo1' -2 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 100 0.01
'TTCL' 'cod2_clo1' 56 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' 'cod2_rel1' 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'VEL' 'cod2_clo1' -0.1 8 1 NA=1 0 0
##
```

Figure 3. TVbutte.O

```
%'OSC_ID' NatFreq m,n escap amp_init phase_init / riseramp plateau fallramp
'v1' 2   1   4   1 NaN/ 10 200 210
'v_rnd1' 2   1   4   1 NaN/ 10 200 210
'ons1_clo1' 2   1   4   1 NaN/ 5 60 65
'ons1_rel1' 2   1   4   1 NaN/ 5 20 25
'ons2_nar1' 2   1   4   1 NaN/ 5 60 65
'cod2_clo1' 2   1   4   1 NaN/ 5 55 60
'cod1_rel1' 2   1   4   1 NaN/ 5 20 25
'cod2_rel1' 2   1   4   1 NaN/ 5 20 25
'cod1_nar1' 2   1   4   1 NaN/ 5 55 60
'cod2_h1' 2   1   4   1 NaN/ 5 55 60

/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
'ons1_clo1' 'ons2_nar1' 1  1 90
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo1' 'v1' 1 1 0
'ons2_nar1' 'v1' 1 1 0
'v_rnd1' 'v1' 1 1 0
'cod1_nar1' 'cod2_clo1' 1  1 45
'cod1_nar1' 'cod1_rel1' 1  1 60
'cod2_clo1' 'cod2_rel1' 1  1 60
'cod2_clo1' 'cod2_h1' 1  1 20
'v1' 'cod1 nar1' 1   1 180
```

Figure 4. PHbutte.O

Figure 4 illustrates PHbutte.O file. This file is divided into two sub-sections: the parameters of timing oscillators (line 1 – 11) and coupling specifications (line 13 – 25). Each line of oscillator parameters consists of the following information: 'OSC_ID' NatFreq m:n escap amp_init phase_init / riseramp plateau fallramp. In line 6, for example, Osc_ID is ‘ons2_nar1’ which is identical to an oscillator label in the TVbutte.O file. NatFreq refers to the natural frequency of a limit cycle oscillator, which is set to 2 Hz for all oscillators in the example file. This natural frequency parameter is used to elicit temporal variation (see Section 2.3.2). m:n refers to the ratio of the natural frequency of oscillator pairs which determines the generalized relative phase. In the example file, m:n is set to 1 since the natural frequency of oscillators corresponding to vowel and consonant gestures is equal to 2 Hz. escap refers to the oscillator escapement, which is set to 4. amp_init is the amplitude at time t0, which is set to 1. phase-init refers to the oscillator phase at time t0. In the

example file, phase-init is set to NaN which means that a random phase is chosen. riseramp, plateau, and fallramp refer to activation and de-activation phases. For example, in line 4, 5 degrees for riseramp mean that the activation of the gesture is started from a value of 0 at 0 degrees to a maximum value of 1 at 5 degrees. 60 degrees for plateau indicate that the gesture stays at the maximum level until the phase reaches 60 degrees. 65 degrees for fallramp means that the activation of gesture goes down to reach a value of 0 again at 65 degrees.

Each line of coupling specifications consists of the following information: 'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase. 'OSC_ID1' and 'OSC_ID2' refer to a pair of oscillator labels. strength1(to OSC1) specifies the relative coupling strength from osc2 onto osc1, and strength2(to OSC2) specifies the coupling strength from osc1 onto osc2. TargetRelPhase refers to a target relative phase for the two oscillators. For example, line 18 shows the coupling specifications between 'ons1_clo1' (Onset C1) and 'v1' (vowel), and their relative coupling strength is equal to 1. Their target relative phase is 0 degrees, which indicates the in-phase relation between the onset and the vowel (c.f., line 25 shows anti-phase relation between the vowel and the coda, specified as 180 degrees; see e.g., Goldstein et al. 2006; Nam et al. 2009).

Based on the TV<id>.O and PH<id>.O files, a gestural score can be computed through TV computation (by clicking the [TV] button in TADA), as shown in Figure 5. The gestural score can be saved as a TV<id>.G file, which contains a gestural score with timing information for each gesture. The gestural score and output time functions can also be saved in .mat file format. This .mat file can be visualized in MVIEW and gestural landmarks can also be parsed with reference to the velocity signal using the findgest function in MVIEW (Tiede 2005). This makes the comparison possible between the simulations from computational modeling and the results from the EMA recordings.

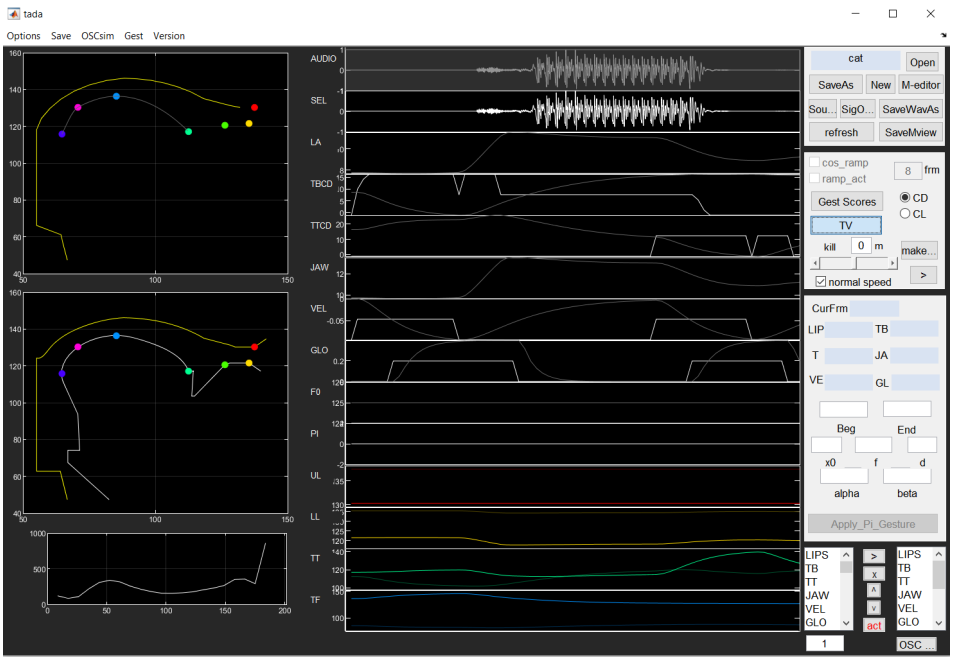


Figure 5. Gestural score in TADA

2.3 Temporal variation in TADA simulations

The goal of the current study is to identify a gestural model that can represent underlying and assimilatory palatalization in Russian. To this end, multiple gestural models are evaluated by comparing their simulation outputs with results from EMA recordings. However, because the analytical approach adopted in Oh et al. (2024) relies on natural temporal variation to reveal differences in gestural coordination, it is necessary for the simulations to incorporate comparable temporal variability. This section therefore examines two methods for eliciting temporal variation in TADA simulations.

2.3.1 Modulating stiffness

Stiffness specifies the intrinsic temporal aspect of each gesture, shaping how quickly

a constriction is formed and released. Because of this property, one potential method for eliciting temporal variation in the duration of the first gesture (G1) is to manipulate gestural stiffness by modulating the freq parameter in the TV<id>.O file. As described in Section 4.1, the freq parameter specifies gestural stiffness, which is typically set to 8 Hz for consonantal gestures and 4 Hz for vocalic gestures. To assess the effectiveness of this approach, the stiffness of /b/ was modulated in the English word /bjut/ 'butte', which served as a test item, and the resulting temporal coordination was examined. Table 1 summarizes the variation in G1 duration for /b/ in 'butte' resulting from changes in stiffness ranging from 3 to 20 Hz.

Table 1. Stiffness and G1 duration

stiffness	G1 duration	stiffness	G1 duration
3	155	12	140
4	155	13	135
5	155	14	135
6	155	15	130
7	150	16	130
8	145	17	130
9	145	18	125
10	140	19	125
11	140	20	125

Because the /bj/ sequence in 'butte' constitutes a clear case of a segmental sequence, this manipulation would be expected to yield a positive correlation between G1 duration and onset-to-onset lag. Contrary to this expectation, however, Figure 6 shows that G1 duration remains independent of onset-to-onset lag when stiffness is manipulated to induce temporal variation. Moreover, given that natural variation in G1 duration observed in the EMA data from Oh et al. (2024) ranges approximately from 100 ms to 500 ms, stiffness modulation fails to generate sufficient variability for meaningful comparison between simulated outputs and empirical results. In sum, manipulating gestural stiffness neither produces adequate temporal variation nor successfully captures the temporal coordination patterns characteristic of segmental sequences in English. The following section therefore considers an alternative method for introducing temporal variation in TADA simulations.

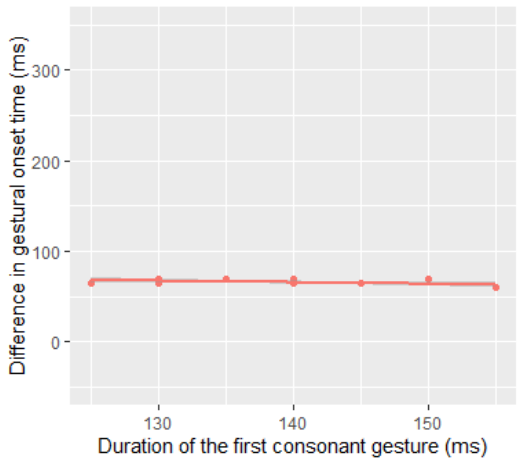


Figure 6. A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis) for stiffness modulation

2.3.2 Modulating natural frequency

Natural frequency (NatFreq), by contrast, refers to the oscillatory rate of the planning units in the coupled oscillator model. Modulating this parameter changes the overall temporal scale of gestural execution, thereby influencing the duration of gestures. Building on this property, another method to elicit temporal variation in TADA simulations is to manipulate NatFreq in the PH<id>.O file. As discussed in Section 2.1., NatFreq refers to the natural frequency of a limit cycle oscillator, which is set to 2 Hz for all oscillators as a default. To test the effectiveness of this method, we modulate the natural frequency of all oscillators in the English word /bjut/ 'butte' as a test word. Table 2 shows the variation in G1 duration for /b/ in 'butte' which is derived by the changes in natural frequency from 0.5 to 3 Hz.

Table 2. Natural frequency and G1 duration

NatFreq	G1 duration	NatFreq	G1 duration	NatFreq	G1 duration
0.5	430	1.4	190	2.3	135
0.6	370	1.5	190	2.4	135
0.7	330	1.6	175	2.5	135
0.8	285	1.7	170	2.6	130
0.9	260	1.8	165	2.7	125

1	240	1.9	150	2.8	125
1.1	225	2	145	2.9	125
1.2	210	2.1	145	3	120
1.3	195	2.2	140		

In contrast to the results reported in Section 2.2.1, modulation of natural frequency generates sufficient temporal variation in G1 duration, ranging from approximately 120 ms to 430 ms, allowing for meaningful comparison between simulated outputs and the EMA data. More importantly, as shown in Figure 7, variation in G1 duration induced by natural frequency modulation is correlated with onset-to-onset lag, consistent with the predictions of the segment sequence hypothesis. Accordingly, natural frequency modulation was adopted as the method for introducing temporal variation in the simulations. For each gestural model, the natural frequency of all oscillators was varied from 0.5 to 3 Hz, and the resulting simulations were compared with the EMA findings. The following section introduces the four gestural models examined in this study.

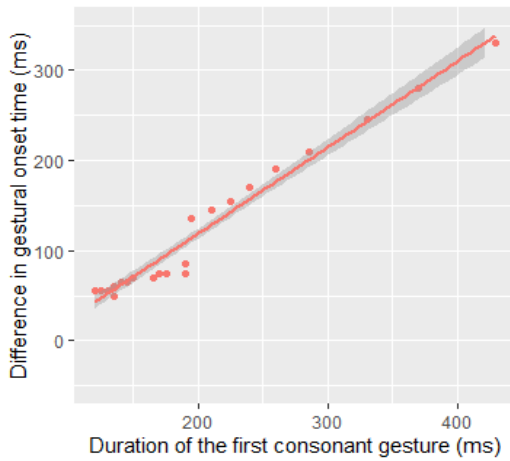


Figure 7. A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis) – Natural frequency modulation

3. Gestural models for underlying and assimilatory palatalization

To identify gestural models capable of representing underlying and assimilatory palatalization in Russian, a pair of target words used in the EMA study—/bʲust/ and /bjut/—was selected. Four gestural models are proposed, each designed to potentially replicate the articulatory patterns observed in the EMA data.

3.1 Model 1

As discussed in 1, the observed differences in tongue body position across palatalization types might be attributable to the blending of the palatal gesture and the following vowel gesture (see also Oh et al. 2024). Model 1 tests whether gestural interaction between /j/ and /u/ alone can account for the articulatory characteristics of assimilatory palatalization, specifically a retracted tongue position at the onset of the tongue body (TB) gesture, a longer onset-to-onset lag, and the absence of an effect of consonant duration on onset-to-onset lag. However, because both underlying and assimilatory palatalization involve the same following back vowel, yet tongue body retraction is observed only in the assimilatory case, additional structural differences are likely required to explain the incomplete neutralization.

In Model 1, the two palatalization types are assumed to differ in their temporal coordination patterns. Gestural representations for /bʲust/ and /bjut/ are illustrated in Figure 8. The primary distinction between the two forms lies in the phasing relationship between the onset consonant gestures. In /bʲust/, the gestures for /b/ and /j/ are coupled in-phase, whereas in /bjut/ they are coupled with a 90-degree phase. For /bjut/, the vowel gesture is coupled in-phase with both onset consonant gestures. As a result of the c-center effect (see, e.g., Shaw et al. 2011), the vowel gesture for /u/ precedes the palatal gesture for /j/. This temporal configuration allows gestural blending between tongue backing associated with /u/ and tongue fronting associated with /j/, which may yield a retracted tongue position at the onset of the TB gesture and a delayed onset-to-onset lag in assimilatory palatalization.

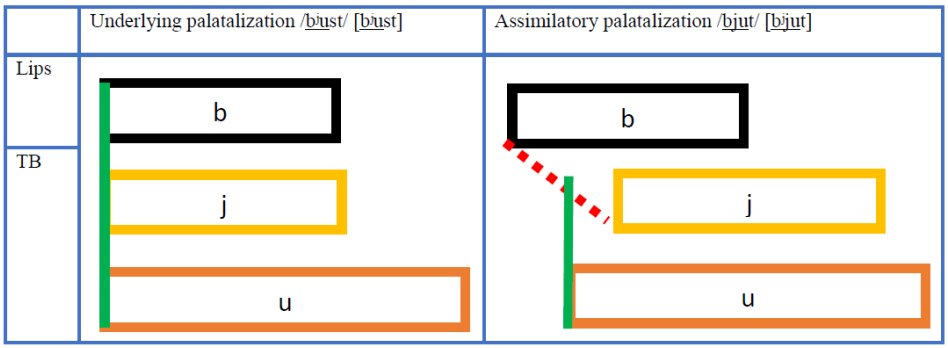


Figure 8. Gestural models for /bʲust/ and /bjut/. The green solid lines indicate in-phase coupling, and the red dotted line indicates 90-degree phase coupling.

Gestural scores for /bʲust/ and /bjut/ were constructed as follows. First, coupling graphs were created via the GEST menu in TADA using ARPABET notation: (B Y-UW_S T) for /bʲust/ and (B Y-UW_T) for /bjut/. The ARPABET symbols and their corresponding IPA representations are provided in Table 3. Identical gestural specifications were applied to both forms, as summarized in Table 4. For /bʲust/, the phasing relationship between /b/ and /j/ was modified to in-phase by setting the TargetRelPhase parameter to 0 in the PHbʲust.O file (see line 16 in Figure 9), while the default phasing relation was retained for /bjut/. The resulting coupling specifications are shown in Figures 9 and 10, respectively. Based on these coupling graphs, gestural scores were generated through TV computation in TADA and saved in .mat format.

Table 3. ARPABET

IPA	ARPABET	IPA	ARPABET	IPA	ARPABET	IPA	ARPABET	IPA	ARPABET
/p/	P	/f/	F	/m/	M	/i/	IY	/u/	UW
/b/	B	/v/	V	/n/	N	/ɪ/	IH	/ʊ/	UH
/t/	T	/θ/	TH	/ŋ/	NX	/e/	EY	/o/	OW
/d/	D	/ð/	DH	/l/	L	/ɛ/	EH	/ɔ/	AO
/k/	K	/s/	S	/ɹ/	R	/æ/	AE	/a/	AA
/g/	G	/z/	Z	/j/	Y	/ʌ/	AH	/aɪ/	AI
/tʃ/	CH	/ʃ/	SH	/w/	W	/ə/	AX	/aʊ/	AW
/dʒ/	JH	/ʒ/	ZH	/h/	HH	/ɚ/	ER	/ɔɪ/	OY

Table 4. Gestural specifications for /b/, /j/, and /u/

IPA	Organ	OSC_ID	TV	Constrict	Target	Stiff	Blending
/b/	Lips	ons1_clo1	LA	CLO	-2	8	100
	Lips	ons1_rel1	LA	REL	11	8	1
	Velum	ons1_clo1	VEL	CLO	-0.1	8	0
/j/	TB	ons2_nar1	TBCL	PAL	95	8	100
	TB	ons2_nar1	TBCD	NAR	2	8	100
	Lips	ons2_nar1	LA	V	8	8	1
/u/	TB	v1	TBCL	UVU/VEL	125	4	1
	TB	v1	TBCD	V	2	4	1
	Lips	v_rnd1	LP	PRO	12	4	1
	Lips	v_rnd1	LA	NAR	5	4	1

```
/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
'ons1_clo1' 'ons2_nar1' 1  1 0
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo1' 'v1' 1 1 0
'ons2_nar1' 'v1' 1 1 0
'v_rnd1' 'v1' 1 1 0
'cod1_crt1' 'cod2_clo1' 1  1 45
'cod1_crt1' 'cod1_rel1' 1  1 60
'cod2_clo1' 'cod2_rel1' 1  1 60
'cod1_crt1' 'cod1_h1' 1  1 20
'v1' 'cod1_crt1' 1   1 180
```

Figure 9. Coupling specifications for /bʲust/

```
/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
'ons1_clo1' 'ons2_nar1' 1  1 90
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo1' 'v1' 1 1 0
'ons2_nar1' 'v1' 1 1 0
'v_rnd1' 'v1' 1 1 0
'cod1_clo1' 'cod1_rel1' 1  1 60
'cod1_clo1' 'cod1_h1' 1  1 20
'v1' 'cod1_clo1' 1   1 180
```

Figure 10. Coupling specifications for /bjut/

3.2 Model 2

Model 2 is proposed to test the prediction that the incomplete neutralization observed in Russian palatalization patterns arises from gestural blending between a velar gesture and a palatal gesture. In this model, the 90-degree phasing relationship between /b/ and /j/ adopted in Model 1 is retained, and an additional velar gesture is introduced, coordinated in-phase with the labial gesture. Model 2 thus differs from Model 1 in that it includes an explicit gesture corresponding to secondary velarization in the assimilatory palatalization condition. Gestural representations for /bʲust/ and /bɯjut/_90-degree-phase are shown in Figure 11. Under this configuration, gestural blending among tongue backing for /ɯ/ and /u/ and tongue fronting for /j/ is expected to yield a more retracted tongue position at the onset of the tongue body (TB) gesture, along with a delayed onset-to-onset lag for assimilatory palatalization.

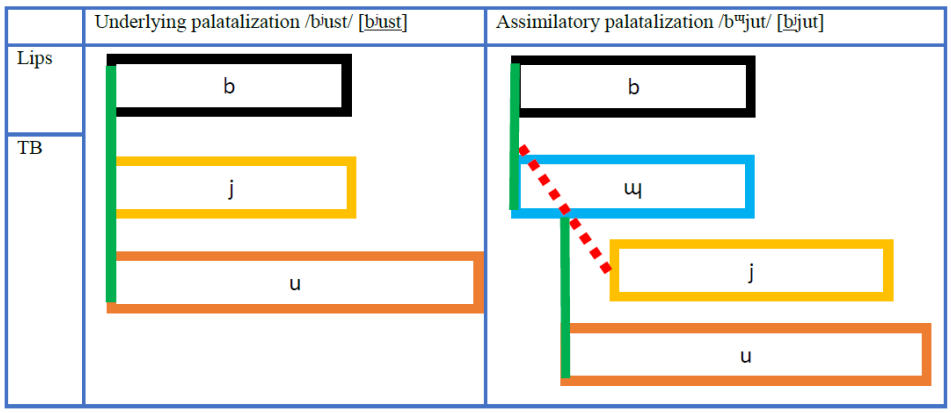


Figure 11. Gestural models for /bʲust/ and /bɯjut/_90-degree-phase. The green solid lines indicate in-phase coupling, and the red dotted line indicates 90-degree-phase coupling.

```
12  /coupling/
13
14  %'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
15  'ons1_clo1' 'ons2_nar1' 1  1 0
16  'ons1_clo1' 'ons3_nar1' 1  1 90
17  'ons1_clo1' 'ons1_rel1' 1  1 65
18  'ons1_clo1' 'v1' 1 1 0
19  'ons2_nar1' 'v1' 1 1 0
20  'ons3_nar1' 'v1' 1 1 0
21  'v_rnd1' 'v1' 1 1 0
22  'cod1_clo1' 'cod1_rel1' 1  1 60
23  'cod1_clo1' 'cod1_h1' 1  1 20
24  'v1' 'cod1_clo1' 1   1 180
```

Figure 12. Coupling specifications for /bɰjut/_90-degree-phase

The gestural score for /bɰjut/_90-degree-phase was constructed as follows. First, a coupling graph was created via the GEST menu in TADA using ARPABET notation (B W Y-UW_T). The velar approximant /ɰ/ was derived from the gestural specifications of /w/ by removing labial gesture specifications in the TVbɰjut_anti.O file. The resulting gestural specifications for /ɰ/ are summarized in Table 5 (see Table 4 for the specifications of /b/, /j/, and /u/). The phasing relationship between /b/ and /ɰ/ was then modified to in-phase by setting the TargetRelPhase parameter to 0 in the PHbɰjut_anti.O file (see line 15 in Figure 12). Gestural and coupling specifications for /bʲust/ follow those described in Section 3.1.

Table 5. Gestural specifications for /ɰ/

IPA	Organ	OSC_ID	TV	Constrict	Target	Stiff	Blending
/ɰ/	TB	ons2_nar1	TBCL	UVU/VEL	125	8	10
	TB	ons2_nar1	TBCD	NAR	2	8	100

3.3 Model 3

Model 3 is similar to Model 2 except that gestures for /b/, /ɰ/, and /j/ are all coupled in-phase with each other for the assimilatory palatalization. Gestural models for /bʲust/ and /bɰjut/_in-phase are visualized in Figure 13. The gestural score for /bɰjut/_in-phase is created in the following way. First, I take the coupling graph for /bɰjut/_90-degree-phase and modify the phasing relation between /b/ and /j/ to be in-phase by changing TargetRelPhase to 0 in the PHbɰjut_in.O file (see line 16

in Figure 14). The gestural specifications for /bɯjut/_in-phase are the same as /bɯjut/_90-degree-phase (see Table 4 and 5). Also, see Section 3.1 for the gestural and coupling specifications for /bʲust/.

Model 3 builds on Model 2 by further modifying the temporal coordination among the consonantal gestures. Specifically, the gestures for /b/, /ɯ/, and /j/ are all coupled in-phase in the assimilatory palatalization condition. Gestural models for /bʲust/ and /bɯjut/_in-phase are illustrated in Figure 13. The gestural score for /bɯjut/_in-phase was obtained by taking the coupling graph from /bɯjut/_90-degree-phase and changing the phasing relationship between /b/ and /j/ to in-phase by setting TargetRelPhase to 0 in the PHbɯjut_in.O file (see line 16 in Figure 14). Gestural specifications remain identical to those used in /bɯjut/_90-degree-phase (see Tables 4 and 5), and the specifications for /bʲust/ are again as described in Section 3.1.

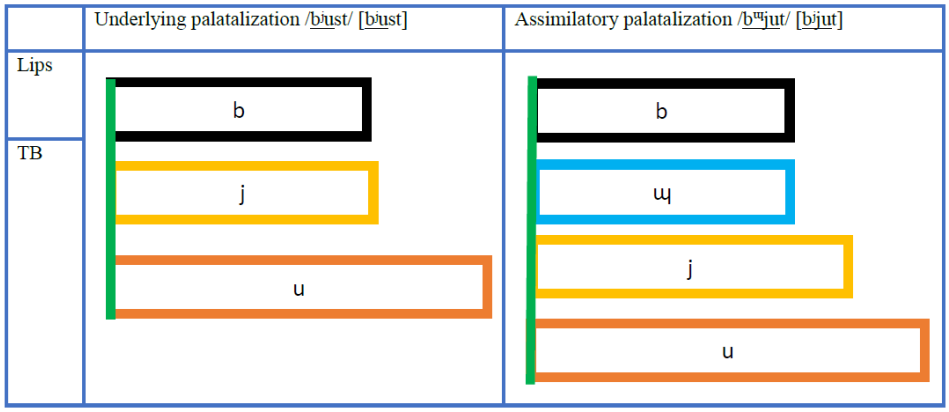


Figure 13. Gestural models for /bʲust/ and /bɯjut/_in-phase. The green solid lines indicate in-phase coupling.

```
/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
'ons1_clo1' 'ons2_nar1' 1  1 0
'ons1_clo1' 'ons3_nar1' 1  1 0
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo1' 'v1' 1 1 0
'ons2_nar1' 'v1' 1 1 0
'ons3_nar1' 'v1' 1 1 0
'v_rnd1' 'v1' 1 1 0
'cod1_clo1' 'cod1_rel1' 1  1 60
'cod1_clo1' 'cod1_h1' 1  1 20
'v1' 'cod1_clo1' 1   1 180
```

Figure 14. Coupling specifications for /bɰjut/_in-phase

3.4 Model 4

As discussed in Section 1, the velar gesture in assimilatory palatalization may begin prior to the palatal gesture for /j/ and overlap with it temporally (see also Oh et al. 2024). Model 4 is designed to test this possibility by introducing eccentric timing between /ɰ/ and /j/, as schematized in Figure 15. The gestural score for /bɰjut/_eccentric was derived from the coupling graph of /bɰjut/_in-phase by modifying the phasing relationship between /ɰ/ and /j/ to an eccentric configuration, implemented by changing TargetRelPhase from 0 degrees to 45 degrees in the PHbɰjut_eccentric .O file (see line 15 in Figure 16). Gestural specifications for /bɰjut/_eccentric are otherwise identical to those used in /bɰjut/_in-phase (see Tables 4 and 5), and the gestural and coupling specifications for /bʲust/ follow Section 3.1.

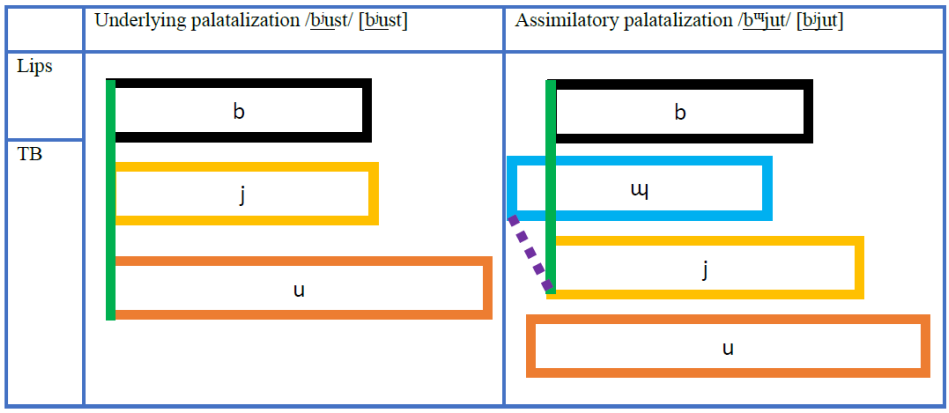


Figure 15. Gestural models for /bʲust/ and /bɯjut/_eccentric. The green solid lines indicate in-phase coupling, and the purple dotted line indicates eccentric timing.

```
/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2) TargetRelPhase
'ons2_nar1' 'ons3_nar1' 1  1 45
'ons1_clo1' 'ons3_nar1' 1  1 0
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo1' 'v1' 1 1 0
'ons2_nar1' 'v1' 1 1 0
'ons3_nar1' 'v1' 1 1 0
'v_rnd1' 'v1' 1 1 0
'cod1_clo1' 'cod1_rel1' 1  1 60
'cod1_clo1' 'cod1_h1' 1  1 20
'v1' 'cod1_clo1' 1   1 180
```

Figure 16. Coupling specifications for /bɯjut/_eccentric

4. Results from TADA simulations

A central finding of Oh et al. (2024) is that the palatal–plain contrast in this context is neutralized, but that the neutralization is phonetically incomplete. Both underlying and assimilatory palatalization exhibit temporal coordination characteristic of complex segments, as reflected in a nearly flat relationship between consonant duration and onset-to-onset lag. This pattern suggests that plain consonants in the assimilatory

palatalization context are realized as palatalized. At the same time, residual effects of an underlying tongue dorsum retraction are observed in assimilatory palatalization: the tongue body (TB) is more retracted at the onset of the palatal gesture, and the onset-to-onset lag is longer, relative to underlying palatalization. The present section evaluates the gestural models introduced in Section 3 by comparing simulation outputs from each model against the EMA results reported in Oh et al. (2024).

4.1 Model 1

As outlined in Section 3.1, Model 1 tests whether gestural blending between /j/ and /u/ alone can give rise to incomplete neutralization in Russian palatalization. In this model, /b/ and /j/ are coupled in-phase in /bʲust/ (underlying palatalization), whereas they are coupled with a 90-degree phase in /bjut/ (assimilatory palatalization), as illustrated in Figure 8.

Figure 17 plots the relationship the correlation between G1 duration (x-axis) and onset-to-onset lag (y-axis) across Status (palatalization types), with least-squares regression lines fit to each condition. For /bʲust/, the regression line is nearly flat, as expected for complex segment coordination. In contrast, simulations for assimilatory palatalization pattern like a segment sequence: onset-to-onset lag increases with G1 duration, yielding a positive correlation. Thus, Model 1 fails to capture the neutralization of temporal coordination observed in the EMA data. The left panel of Figure 18 shows the horizontal (front–back) position of the TB at the onset of the palatal gesture. Although the TB is slightly more retracted for assimilatory palatalization than for underlying palatalization, the magnitude of this difference is minimal and substantially smaller than the approximately 1.5 mm difference reported in the EMA results from Oh et al. (2024). The right panel of Figure 18 further shows that onset-to-onset lag is longer for assimilatory than for underlying palatalization. In sum, Model 1 yields a delayed onset-to-onset lag for assimilatory palatalization but fails to reproduce the complex-segment temporal coordination shared by both palatalization types in the EMA data, as well as the robust spatial retraction of the TB. Results from Model 2 are presented next.

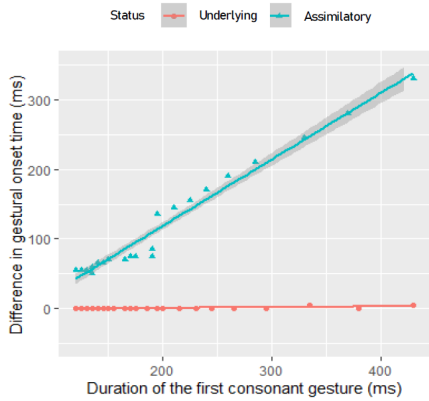


Figure 17. Model 1 – A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis)

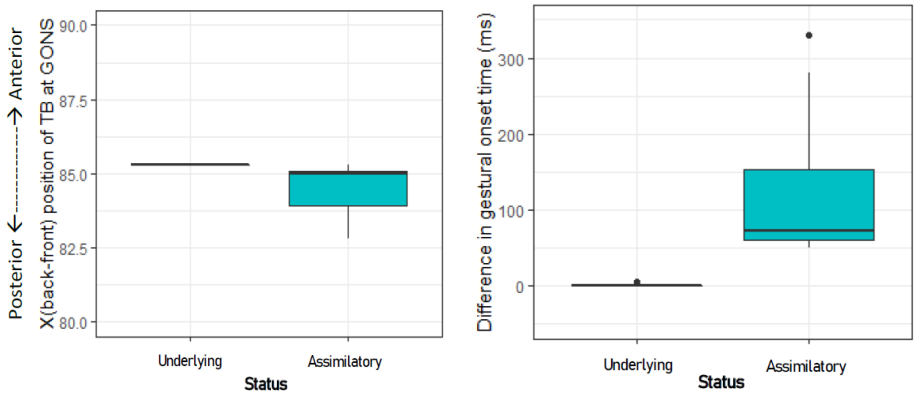


Figure 18. Model 1 – A boxplot of TB position (mm) at palatal gesture onset (left), and onset-to-onset lag across Status (right)

4.2 Model 2

Model 2 extends Model 1 by introducing a gesture for secondary velarization in assimilatory palatalization (see Section 3.2 and Figure 11). Under this model, gestures for /b/ and /j/ are coupled in-phase with each other in /bʲust/ (underlying palatalization), while gestures for /b/ and /j/ are coupled 90 degree phase with each other and gestures for /b/ and /ɯ/ are coupled in-phase in /bɯjut/_90-degree-phase (assimilatory palatalization).

As shown in Figure 19, simulations for /bʲust/ (underlying palatalization) again exhibit a nearly flat regression line, consistent with complex-segment coordination. However, /bɯjut/_90-degree-phase (assimilatory palatalization) continues to show a positive correlation between G1 duration and onset-to-onset lag, indicating segment-sequence–like behavior, similar to Model 1. Unlike Model 1, however, Model 2 produces a more retracted TB position for assimilatory palatalization at the onset of the palatal gesture, as shown in the left panel of Figure 20. The lag between the gesture onsets is longer for the assimilatory palatalization than for the underlying palatalization (see Figure 20, right panel).

Thus, while Model 2 successfully captures the spatial asymmetry between palatalization types, it fails to reproduce the temporal coordination pattern observed for assimilatory palatalization in the EMA data. Results from Model 3 are discussed in the following section.

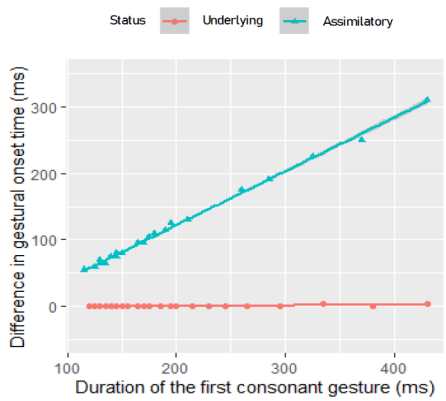


Figure 19. Model 2 – A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis)

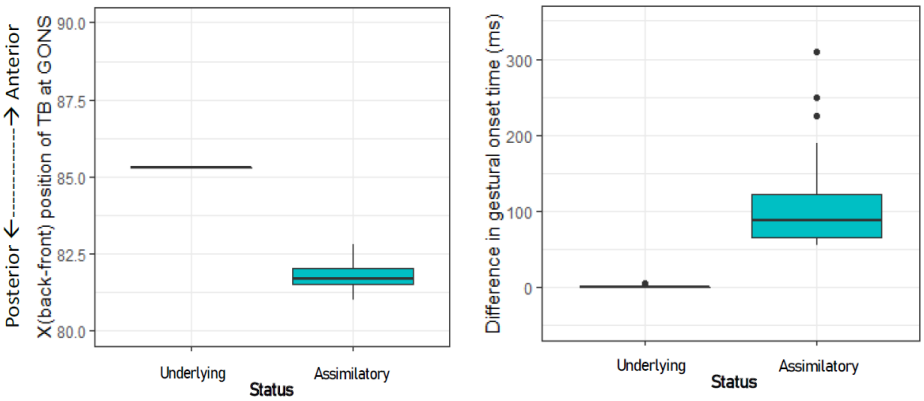


Figure 20. Model 2 – A boxplot of TB position (mm) at palatal gesture onset (left), and onset-to-onset lag across Status (right)

4.3 Model 3

Model 3 examines whether in-phase coupling among /b/, /ɯ/, and /j/, together with gestural blending between /ɯ/ and /j/, can account for incomplete neutralization (see Section 3.3). The gestural configurations for /bʲust/ and /bɯjut/_in-phase are shown in Figure 13.

As illustrated in Figure 21, both underlying (/bʲust/) and assimilatory palatalization (/bɯjut/_in-phase) yield nearly flat regression lines, indicating complex-segment coordination in both cases. This pattern closely matches the temporal coordination observed in the EMA recordings. However, Model 3 fails to produce differences between palatalization types in either the spatial position of the TB or onset-to-onset lag (Figures 22 left and right panels, respectively).

In summary, Model 3 successfully reproduces the shared temporal coordination of underlying and assimilatory palatalization but does not capture the spatial retraction or delayed onset associated with assimilatory palatalization. The results of Model 4 are presented next.

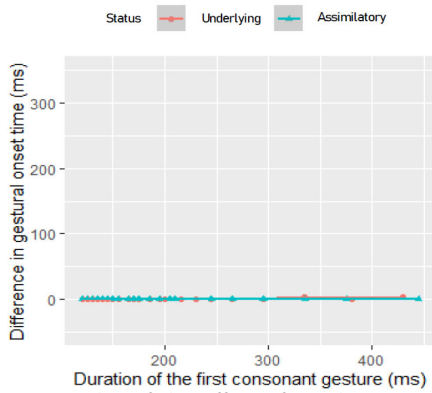


Figure 21. Model 3 – A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis)

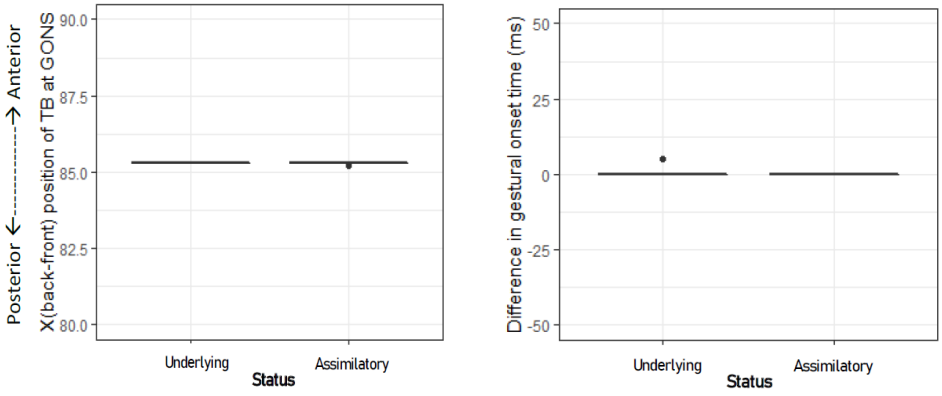


Figure 22. Model 3 – A boxplot of TB position (mm) at palatal gesture onset (left), and onset-to-onset lag across Status (right)

4.4 Model 4

Model 4 introduces eccentric timing between /ɯ/ and /j/ in assimilatory palatalization, as schematized in Figure 15. Specifically, In particular, gestures for /ɯ / and /j/ are coupled eccentric-phase (45 degrees) with each other, and gestures for /b/ and /j/ are coupled in-phase in /bɯjut/_eccentric (assimilatory palatalization).

As shown in Figure 23, simulations for both underlying and assimilatory palatalization exhibit nearly flat regression lines, with only a slight upward trend for

assimilatory palatalization (/bɯjut/_eccentric). This pattern indicates complex-segment coordination comparable to that observed in the EMA data. In addition, Model 4 produces a more retracted TB position for assimilatory palatalization at the onset of the palatal gesture (see Figure 24a). The lag between the gesture onsets is also longer for the assimilatory palatalization (/bɯjut/_eccentric) than for the underlying palatalization (/bʲust/) as shown in Figure 24b.

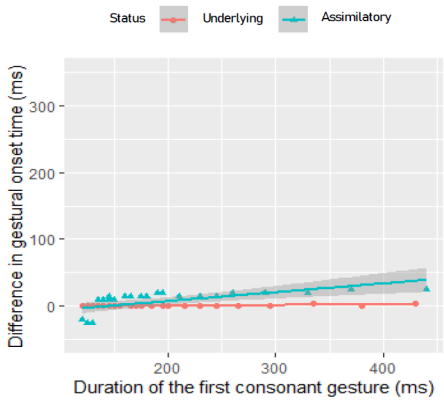


Figure 23. Model 4 – A scatter plot of the effect of G1 duration (x-axis) on onset-to-onset lag (y-axis)

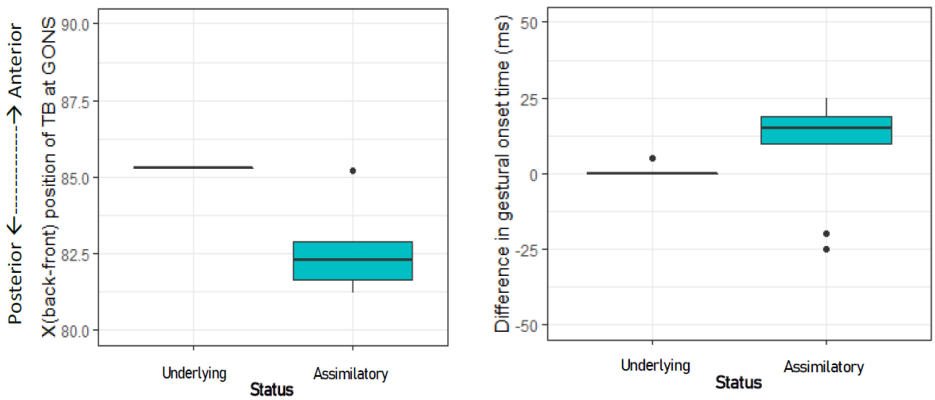


Figure 24. Model 4 – A boxplot of TB position (mm) at palatal gesture onset (left), and onset-to-onset lag across Status (right)

Taken together, the results demonstrate that among the four models evaluated, only Model 4 successfully reproduces the full set of EMA findings: complex-segment temporal coordination for both palatalization types, coupled with a more retracted tongue body position and a delayed onset-to-onset lag for assimilatory palatalization. Models 1 and 2 fail to capture the observed temporal coordination, while Model 3 fails to derive the spatial and temporal asymmetries between palatalization types.

5. Discussion

5.1 Overview

The aim of this study was to identify the precise internal gestural organization that can account for incomplete neutralization between underlying and assimilatory palatalization in Russian. To this end, the current study systematically explored how different internal gestural configurations can give rise to temporal coordination patterns and subtle articulatory differences observed in Oh et al (2024). Specifically, four gestural models were implemented in the articulatory-based synthesizer Task Dynamic Application (TADA), and the resulting simulation outputs from each model were directly compared with the EMA findings reported in the previous work.

Model 1 tested whether gestural blending between the palatal glide /j/ and the following vowel /u/ alone could give rise to incomplete neutralization. The remaining models incorporated both palatal and velar gestures in the assimilatory palatalization condition, differing in their temporal coordination. Models 2 and 3 examined the hypothesis that incomplete neutralization arises from gestural blending between velar and palatal gestures. The only difference between Model 2 and Model 3 was that the palatal and velar gestures were coordinated anti-phase with each other in Model 2 and coordinated in-phase with each other in Model 3. Model 4, by contrast, evaluated whether an earlier onset of the velar gesture relative to the palatal gesture, realized through eccentric temporal coordination, could derive the observed patterns.

Across models, only Model 4 successfully reproduced the full set of EMA findings. The failure of Models 1–3 to replicate the EMA outcomes can be traced to their assumptions about intergestural coupling. Models 1 and 2 posit anti-phase

coordination between /b/ and /j/, which yields segment-sequence timing similar to the English /bj/ sequence reported in Shaw et al. (2021). While these configurations produce a delayed onset-to-onset lag (Model 1) or tongue body retraction (Model 2), they cannot capture the complex-segment temporal coordination characteristic of underlying and assimilatory palatalization in Russian. By contrast, Model 3 assumes in-phase coupling among /b/, /ɯ/, and /j/ in assimilatory palatalization, which successfully generates complex-segment timing but fails to derive the robust tongue body retraction at the palatal onset observed in the EMA data. In contrast, Model 4 yielded complex-segment coordination for both underlying and assimilatory palatalization, together with a more retracted tongue body position at the onset of the palatal gesture and a longer onset-to-onset lag in the assimilatory condition. To facilitate comparison across models, Table 6 summarizes the simulation outcomes in terms of temporal coordination, tongue body retraction, and onset-to-onset lag. This finding extends previous articulatory phonology accounts by showing that internal gestural coordination can play a crucial role in shaping surface phonetic patterns, including cases of incomplete neutralization.

Table 6. Simulation outcomes across gestural models

	Temporal Coordination	TB Retraction	Onset-to-Onset Lag
Model 1	X	X	O
Model 2	X	O	X
Model 3	O	X	X
Model 4	O	O	O

5.2 The source of velar gesture in assimilatory palatalization

As summarized above, Model 4, which requires an earlier onset of the velar gesture relative to the palatal gesture, successfully replicated the EMA data. This naturally raises the question of where the velar gesture originates: is it simply an artifact of coarticulation with the back vowel /u/, or does it reflect an independent articulatory source? The most intuitive explanation for the velar gesture would be simple coarticulation between /j/ and /u/, but this possibility was already tested. In fact, both the in-phase coordination of /j/ and /u/ posited for underlying palatalization and the anti-phase coordination posited for assimilatory palatalization in Model 1 failed to

reproduce the observed EMA patterns. This indicates that the velar gesture cannot be attributed to vowel-driven blending alone.

A more plausible interpretation is that the velar gesture comes from the "plain" consonants. Articulatory studies on "plain" consonants in Russian have shown that the plain consonants are accompanied by a secondary dorsal gesture (Litvin 2014; Roon and Whalen 2019). For instance, Roon and Whalen (2019) demonstrated that Russian 'plain' consonants consistently exhibit dorsal gestures across manners of articulation and syllable positions (initial vs. final) with the constriction location ranging from velar to uvular across speakers. This phonetic observation is consistent with phonological evidence such as the /i/-backing process, where the front vowel /i/ is realized as central [ɨ] after "plain" consonants, indicating a secondary velarization that interacts with adjacent segments (see Padgett 2001, 2003; Litvin 2014). These findings suggest that the word /pjot/ 'drink (3ps pres)' may more accurately represented as /pɯjot/. That is, the velar gesture may be better understood not as an abrupt insertion nor as a simple vowel-driven effect, but as an inherent feature of the Russian "plain" consonants. The question of why this velar gesture surfaces specifically through eccentric timing between /ɯ/ and /j/ will be taken up in the following section.

5.3 Temporal coordination of two secondary articulation gestures

Recent work has shown that some languages exhibit simple onset timing, in which consonants within a cluster are coupled anti-phase with each other, but only the rightmost consonant is coupled in-phase with the vowel (Goldstein et al. 2007; Shaw et al. 2011; Tilsen et al. 2012; Hermes et al. 2017). The gestural organization in Model 4 closely resembles the simple onset timing. That is, like the simple onset timing in which only the rightmost prevocalic consonant is coordinated in-phase with the vowel, Model 4 shows that among the secondary articulations only the palatal gesture is coupled in-phase with the labial gesture. Considering there are two secondary articulations involved in cases of assimilatory palatalization, it is possible that these two secondary articulation gestures are coupled in eccentric-phase with each other, just like onset clusters are anti-phase with each other. Given the presence of multiple secondary articulations, this configuration suggests that secondary gestures may be

temporally organized in a manner analogous to consonantal clusters, exhibiting non-synchronous coordination without disrupting complex-segment timing.

Furthermore, eccentric timing has not only been proposed for Russian but also observed cross-linguistically. For instance, eccentric coordination of CV sequences has been reported in the tonal language Mandarin (Gao 2008) as well as in the non-tonal language Tibetan (Geissler 2021). In addition, Goldstein (2011) noted that "eccentric coupling is used to coordinate consonant gestures in an onset or a coda cluster," highlighting that such timing relations may play a broader role in the gestural organization across languages. Against this broader backdrop, however, we argue that the coexistence of velarization and palatalization in Russian onset clusters specifically motivates the positing of a 45° eccentric phase in assimilatory palatalization. In this light, the 45° eccentric timing among secondary articulations seems more plausibly interpreted as language-specific to Russian rather than as a default coordination setting for secondary articulations.

While simulations with gestural blending between the velar and palatal gestures and eccentric timing showed the same results that were observed in the EMA data, it is also possible that such eccentric timing arises as a by-product of other gestural coordination mechanisms. Another possible scenario is that the eccentric timing between the velar and palatal gestures might be attributable to a competitive coupling between the two gestures, such that they are both timed in-phase to the labial gesture but anti-phase (180°) to each other. However, simulations implementing competitive coupling revealed no difference in tongue body position between underlying and assimilatory palatalization, and showed a longer onset-to-onset lag for underlying palatalization, which contradicts the EMA findings (see Appendix for details). Moreover, we explored other competitive coupling graphs with four different phasing relations (90°, 65°, 45°, and 20°) for the velar and palatal gestures, but they did not derive the empirical pattern either. The competitive coupling graph with a relative phase of 20° of a gesture's oscillator was shown as an example in Appendix. These results suggest that competitive coupling alone is insufficient to account for incomplete neutralization in Russian palatalization.

These findings highlight that secondary articulations may interact in systematic ways, and that their relative timing can shape language-specific phonetic outcomes. Russian palatalization provides a case study illustrating how complex segmental organization can reveal broader principles of gestural coordination.

5.4 Summary

This study investigated four gestural models of underlying and assimilatory palatalization in Russian using TADA and evaluated their predictions against EMA data. The model that best matched the empirical findings involved eccentric-phase coordination between velar and palatal gestures in the assimilatory palatalization condition, alongside in-phase coupling between the palatal and labial gestures. In contrast, underlying palatalization was modeled without a velar gesture and with in-phase coordination between labial and palatal gestures.

Simulations from this model reproduced complex-segment temporal coordination for both palatalization types, while also capturing the subtle but systematic articulatory differences: greater tongue body retraction at the onset of the palatal gesture and a longer onset-to-onset lag in assimilatory palatalization. These results support an analysis in which incomplete neutralization in Russian arises from the interaction of gestural blending and the eccentric timing between secondary articulations. More broadly, the findings align with previous evidence for secondary velarization or uvularization in Russian plain consonants and provide a computationally explicit account of the EMA results reported in Oh et al. (2024).

References

Browman, Catherine P. and Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3: 219–252.

Browman, Catherine P. and Louis Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45(2-4): 140-155.

Browman, Catherine P. and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6(2): 201–251.

Browman, Catherine P. and Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3–4): 155–180.

Browman, Catherine P. and Louis Goldstein. 1995. Dynamics and articulatory phonology. In Robert F. Port and Timothy van Gelder (eds.), *Mind as motion: Explorations in the dynamics of cognition*, 175–194. Cambridge, MA: MIT Press.

Gao, Man. 2008. *Tonal alignment in Mandarin Chinese: An articulatory phonology account.* Unpublished PhD Dissertation. Yale University.

Geissler, Chris. 2021. *Temporal articulatory stability, phonological variation, and lexical contrast preservation in diaspora Tibetan*. Unpublished PhD Dissertation. Yale University.

Goldstein, Louis. 2011. Back to the past tense in English. In Rodrigo Gutiérrez-Bravo, Line Mikkelsen, and Eric Potsdam (eds.), *Representing language: Essays in honor of Judith Aissen*, 69–88. Linguistics Research Center, UC-Santa Cruz Department of Linguistics.

Goldstein, Louis, Dani Byrd, and Elliot L. Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In Michael A. Arbib (ed.), *Action to language via the mirror neuron system*, 215-249. Cambridge, MA: MIT Press.

Goldstein, Louis, Ioana Chitoran, and Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. In Jürgen Trouvain and William Barry (eds.), *Proceedings of the XVIth International Congress of Phonetic Sciences*, 241-244. Saarbrücken: University of Saarlandes.

Goldstein, Louis, Hosung Nam, Elliot L. Saltzman, and Ioana Chitoran. 2009. Coupled oscillator planning model of speech timing and syllable structure. *Frontiers in phonetics and speech science*, 239-250. Beijing: Commercial.

Hermes, Anne, Doris Mücke, and Bastian Auris. 2017. The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics* 64: 127-144.

Kochetov, Alexei. 2011. Palatalization. In Marc van Oostendorp, Colin Ewen, Beth Hume, and Keren Rice (eds.), *Companion to phonology* 3, 1666–1690. Oxford: Wiley Blackwell.

Kühnert, Barbara, Philip Hoole, and Christine Mooshammer. 2006. Gestural overlap and C-center in selected French consonant clusters. In Hani Camille Yehia, Didier Demolin, and Rafael Laboissiere (eds.), *Proceedings of the 7th International Seminar on Speech Production*, 327–334. Pampulha: Cefala.

Litvin, Natallia. 2014. *An ultrasound investigation of secondary velarization in Russian.* MA Thesis. University of Victoria.

Nam, Hosung, Louis Goldstein, and Elliot L. Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. In Christophe Coupé and Egidio Marsico (eds.), *Approaches to phonological complexity,* 297-328. Berlin: De Gruyter Mouton.

Nam, Hosung, Louis Goldstein, Elliot L. Saltzman, and Dani Byrd. 2004. TADA: An enhanced, portable Task Dynamics model in MATLAB. *The Journal of the Acoustical Society of America* 115(5): 2430.

Nam, Hosung, Louis Goldstein, Catherine P. Browman, Philip Rubin, Michael Proctor, and Elliot L. Saltzman. 2006. *TADA (task dynamics application) manual.* New Haven, CT: Haskins Laboratories.

Nam, Hosung, Vikramjit Mitra, Mark K. Tiede, Mark Hasegawa-Johnson, Carol Espy-Wilson, Elliot L. Saltzman, and Louis Goldstein. 2012. A procedure for estimating gestural scores from speech acoustics. *The Journal of the Acoustical Society of America* 132(6): 3980-3989.

Oh, Sejin, Jason A. Shaw, Karthik Durvasula, and Alexei Kochetov. 2024. Russian assimilatory palatalization is incomplete neutralization. *Laboratory Phonology* 15(1): 1-37.

Padgett, Jaye. 2001. Contrast dispersion and Russian palatalization. In Elizabeth Hume and Keith Johnson (eds.), *The role of speech perception in phonology,* 187–218. San Diego, CA: Academic Press.

Padgett, Jaye. 2003. The emergence of contrastive palatalization in Russian. In Eric Holt (ed.), *Optimality Theory and language change,* 307–335. Dordrecht: Kluwer.

Pouplier, Marianne. 2020. Articulatory phonology. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics.* Oxford: Oxford University Press.

Roon, Kevin D. and Douglas H. Whalen. 2019. Velarization of Russian labial consonants. In Sasha Calhoun, Paola Escudero, Marija Tabain, and Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 3488–3492. Canberra: Australasian Speech Science and Technology Association Inc.

Saltzman, Elliot L. and Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1(4): 333-382.

Saltzman, Elliot L., Hosung Nam, Jelena Krivokapić, and Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In Plínio Barbosa, Sandra Madureira, and César Reis (eds.), *Proceedings of the 4th International Conference on Speech Prosody (Speech Prosody 2008)*, 175-184. Boston, MA: International Chinese Statistical Association (ICSA).

Shaw, Jason A., Adamantios I. Gafos, Philip Hoole, and Chakir Zeroual. 2011. Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology* 28(3): 455-490.

Tiede, Mark K. 2005. *MVIEW: Software for visualization and analysis of concurrently recorded movement data.* New Haven, CT: Haskins Laboratories.

Tilsen, Sam, Draga Zec, Christina Bjorndahl, Becky Butler, Marie-Josee L'Esperance, Alison Fisher, Linda Heimisdottir, Margaret Renwick, and Chelsea Sanker. 2012. A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters. *Cornell Working Papers in Phonetics and Phonology,* 51-81. Ithaca, NY: Cornell University.

Appendix

As discussed in Section 5.3, the eccentric timing between the velar and palatal gestures could alternatively be analyzed as arising from competitive coupling between the two gestures, such that both gestures are coupled in-phase with the labial gesture but anti-phase with each other. Under this configuration, simulations exhibit temporal coordination characteristic of complex segments for both /bʲust/ (underlying palatalization) and /bɯjut/_in-phase (assimilatory palatalization), as evidenced by the

nearly flat regression lines shown in Figure 1 left, with only a slight upward trend for the assimilatory condition. This pattern closely mirrors the temporal coordination observed in the EMA recordings. However, the competitive coupling configuration fails to capture the spatial and temporal asymmetries reported in the articulatory data. As shown in Figure 2 (left), no difference emerges between underlying and assimilatory palatalization in the tongue body position at the onset of the palatal gesture. Moreover, the onset-to-onset lag is longer for the underlying palatalization than for the assimilatory palatalization (Figure 2, right), a pattern that directly contradicts the EMA findings. Comparable results are obtained when the relative phase between the velar and palatal gestures is set to 20° in the competitive coupling configuration (see Figures 1 right and 3).

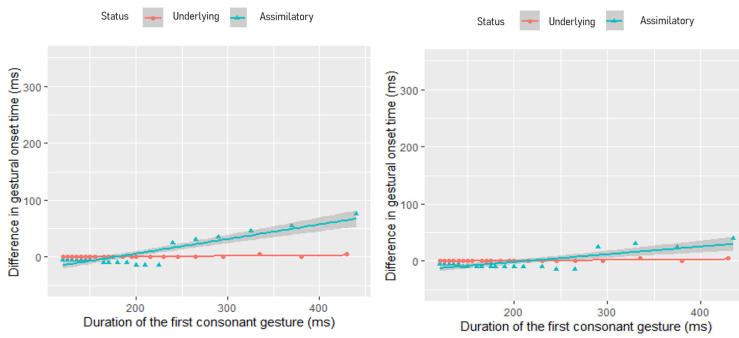


Figure 1. A scatter plot of the effect of G1 duration on onset-to-onset lag - Competitive coupling (180°) (left), and Competitive coupling (20°) (right)

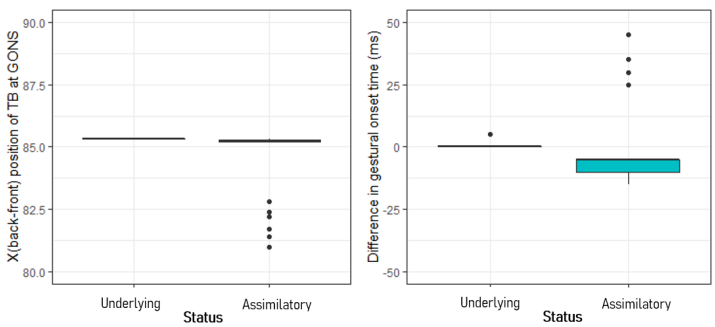


Figure 2. Competitive coupling (180°) – TB position (mm) at palatal gesture onset (left), and onset-to-onset lag (right)

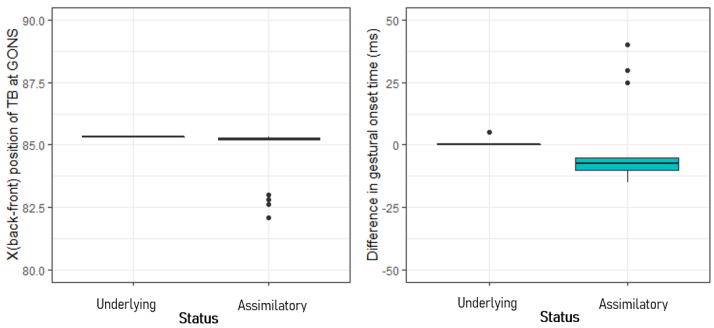


Figure 3. Competitive coupling (20°) – TB position (mm) at palatal gesture onset (left), and onset-to-onset lag (right)

Sejin Oh
Assistant Professor
Department of English Education
Jeju National University
102, Jejudaehak-ro, Jeju-si,
Jeju-do, Republic of Korea
E-mail: sejinoh@jejunu.ac.kr

Received: 2026. 02. 02.
Revised: 2026. 03. 30.
Accepted: 2026. 05. 04.