



Some articulatory reflexes observed in intervocalic consonantal sequences: Evidence from Korean place assimilation*

Minjung Son**

Department of English Language and Literature, Hannam University, Daejeon, Korea

Abstract

This paper examines kinematic characteristics of /pk/ clusters, as compared to /kk/ and /pp/ with varying vowel contexts and speech rate. The results of EMMA data from eight Seoul-Korean speakers indicate as follows. Firstly, comparing /pk/ to /pp/ sequences, lips closing movement was faster and spatially greater in the /a/-to-/a/ context while temporally longer in the /i/-to-/i/ context. It was smaller in spatial displacement and shorter in temporal duration in /pk/ sequences. Peak velocity did not vary. Secondly, comparing /pk/ with /pp/ and /kk/ controls, lip aperture was less constricted in the /a/-to-/a/ context than /i/-to-/i/, but the maximum contact between the upper and lower lips was invariant across different vocalic contexts within /pk/ sequences (/apka/=/ipki/). Categorical reduction of C1 in /pk/ sequences fell in with the low-vowel and fast-rate conditions with across-/within-speaker variability. Gradient reduction of C1 was observed in all C1C2 types, being more frequent in fast rate. Lastly, the jaw articulator was a stable indicator of rate effects. The implication of the current study is that gestural reduction occurs with categorical reduction and general spatiotemporal weakening in the assimilating contexts, while *quantitative* properties of gestures may be a reason for gradient reduction, not necessarily confined to place assimilation.

Keywords: categorical reduction, partial (or gradient) reduction, place assimilation, vowel context effects, rate effects, jaw

1. Introduction

Place assimilation has been widely attested in many languages. In a typological study on the targeting of place assimilation (Jun, 1995, 2004), some languages demonstrate three places of articulation (coronal, labial, and velar) as a target (e.g., Diola Fogy, Malay, Thai, Nchufie, and Yoruba), and others only allow coronal (e.g., Catalan, English, German, Toba Batak, and Yakut). With respect to its occurrences, coronal is known to be the most fre-

quently selected target. However, targeting exclusively limited to coronal and labial is typologically very infrequent (Jun, 1995, 2004; see also Browman & Goldstein, 1995; Byrd, 1992; Chen, 2003). Place assimilation has been relatively well examined using various theoretical frameworks. At the outset of generative phonology, feature change was responsible for the phonological process (e.g., [+labial]→[+velar] / __ [+velar] in Kim-Renaud (1974)). However, this traditional approach does not provide an explicit mechanism to explain how a speaker probes into and

* This work was supported by NIH Grant DC-00403 conferred upon Catherine T. Best (PI) and Haskins Laboratories.

** minjungson@hnu.kr, Corresponding author

Received 30 April 2020; Revised 15 June 2020; Accepted 15 June 2020

© Copyright 2020 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

singles out (a set of) specific feature(s) under consideration in his/her phonological system. Later, in feature geometry theory (Clements, 1985), feature specifications were defined in a hierarchical way; for that reason, the target [+labial] and the trigger [+velar] were both governed by the *Place* node; consequently, an adjacent set of the target and trigger is close enough so that the former is directly influenced by the latter.

V-to-C formant transitions are important perceptual cues to identify place of articulation (Borden & Harris, 1984). With respect to anticipatory place assimilation (e.g., C1C2→C2C2), it can be motivated by weak perceptibility inherent in coda (Byrd, 1992; Fujimura et al., 1978; Jun, 2004; Krakow, 1989; Steriade 2000, 2001). Acoustic formant transition is the result of coarticulation called *gestural overlap* (Browman & Goldstein, 1986, 1989, 1990a, 1990b, 1990c, 1992, 1995, among others). With increased gestural overlap between C1 and C2, the formant transitional information of C1 can be gradually replaced by that of C2, all things being equal, which is in turn hypothesized to affect speech perception (Browman & Goldstein, 1990c; Byrd, 1992). Another articulatory factor contributing to place assimilation is gestural reduction in C1, either categorical, gradient, or both (Jun, 2004; Kochetov & Pouplier, 2008; Son, 2008; Son et al., 2007). In Chen's (2003) perceptual recovery algorithm study using the outputs of gestural simulation, she found that complete recovery of voiced coronal stop /d/ in the coda of /d#b/ sequences was less likely to take place with increased (or more) overlap, while recovery of underlying voiced labial stop /b/ in the reversed order (/b#d/) was consistently stronger in the performance of the perceptual recovery algorithm. This means that increased gestural overlap in the assimilating /d(#)/b/ sequences is responsible for a weakened perceptual bias to [b(#)/b]. Based on this, increased (or more) gestural overlap is considered to be responsible for a reduced perceptual bias, which could have subsequently induced speakers' miscopying of /d/ as articulatorily diminished.

Revisiting gestural reduction of C1, Browman & Goldstein (ms.) are somewhat reserved in deciding on what could have caused gestural reduction of C1; rather than giving full credit to perceived reduction of C1 with increased overlap, they are also open to the possibility of 'some independent reason' (p. 16). Meanwhile, the reversed order of labial-coronal (e.g., /p#t/ → [pt], not *[tt], 'pump tires' (Byrd, 1992: 19)) avoids place assimilation. Using the output of synthesized speech as the input of a perceptual experiment with human listeners, coronal /d/ in C1, not labial coda, was perceived as assimilated when C1 and C2 were quite overlapped. Byrd (1992) argued that this might be because the lips gesture in C1 moves slower than the tongue tip gesture; consequently, labial in coda is unlikely to be obscured by the subsequent agile articulation of the tongue tip, and thus less susceptible to an 'acoustically hidden' segment (Browman & Goldstein, 1990a: 304). To quote Browman & Goldstein (1990b: 422), "Gestures, however, have quantitative (gradient) articulatory properties..." According to them, gesture-based input specifications in C1C2 sequences are considered to provide explanatory accounts for place assimilation as attested in American English (e.g., the string of 'perfect *memory*' available in the University of Wisconsin X-ray microbeam data). They observed that a small tongue tip gesture still remained after most of the tongue tip constriction degree had lessened, with increased gestural overlap even in fluent and casual speech.

Korean assimilating sequences in the low-vowel context (/a/-to-/a/) demonstrate inter-speaker variability in an EMMA (Perkell et al., 1992) study (Kochetov & Pouplier, 2008). The tongue tip gesture was either categorically deleted, partially (or gradiently) reduced, or unreduced (e.g., /t(#)/p/ and /t(#)/k/). In particular, partially reduced tokens occurred with inter-speaker variability (e.g., one out of the three speakers), while categorical reduction was observed across speakers. Only one subject from Kochetov & Pouplier (2008) showed categorical as well as partial (or gradient) reduction in the production of coronal /t/: partial (or gradient) reduction was quite limited in occurrence if there was any.

With respect to the nature of gestural reduction in a labial target within /p(#)/k/ sequences using EMMA (Perkell et al., 1992), Korean place assimilation has been characterized by categorical reduction (Son et al., 2007). However, subsequent articulatory studies have shown evidence for rare occurrences of gradient (or partial) reduction in terms of spatiotemporal measurements of the lip aperture gesture such as constriction minima combined with constriction duration (Son, 2008). Gradiently reduced coda /p/ in the assimilating /ap(#)ka/ sequences was more frequent in occurrence (3.8% in the coda of production, being limited to the across-word boundary condition), although there was inter-speaker variability (one out of five speakers), compared to gradiently reduced onset /k/ in C2 (1.3% in the onset of production, including both the within-/across-word boundary conditions). Son (2013) showed gestural reduction in terms of kinematic measurements such as closing acceleration duration, overall closing duration, and constriction duration in the lip aperture gesture (/apka/ (assimilation) < /apta/ (non-assimilation)). Jun (1996) attributed perceptual place assimilation to the incomplete lips closing gesture (i.e., partial reduction in Jun (1996)) for /p/ in the environment of __/k/) based on his aerodynamic study, which can, in principle, fall in anywhere along a continuum with categorical reduction at one extreme and complete closure at the other. In that case, the emergence of categorically reduced /p/ tokens, if not all, could be interpreted as an extreme manifestation of gradient reduction along a continuum. His oral pressure data from fourteen speakers demonstrated reduction of C1 for /VpkV/ sequences in 47% of production in various vocalic contexts (e.g., /ipki/, /upku/, /ipku/, /upki/).

With regard to perceived assimilation, previous studies agree that labial is less susceptible to place assimilation as compared to coronal. This is because more sluggish movement is less likely to be obscured by a following segment with an agile articulatory movement (Browman & Goldstein, 1990a, 1990c). In the tongue musculature of primates, type I fibers are hypothesized to be related to slower movements of the posterior of the tongue body and type IIA fibers to rapid movements of the apex of the tongue (DePaul & Abbs, 1996). In this sense, the movements of the lips can be gradually covered by the sluggish movements of the tongue back in C2. However, in a perceptual study using various types of /pk/ sequences acquired through simultaneous air pressure-acoustic methodology (e.g., non-overlapped, highly overlapped, reduced C1, and control), highly overlapped /pk/ sequences with fully produced C1 were still perceived as unassimilated (Jun, 1996). Likewise, perceptual consequences have also not been manifested through manipulating degrees of

gestural overlap: either: human listeners' perception was not finely sensitive to C1 reduction (Son et al., 2007).

1.1. Research questions

In the current study, we focus on gestural reduction of C1. Partly due to different experimental or analytical methodologies employed for kinematic studies on assimilating contexts and their controls, there has not been a single kinematic study which has provided, to the best of our knowledge, reduction frequency (either categorical, gradient, or both) along with kinematic characteristics during lip aperture closing movement. In the current study, we are concerned with systematically describing some articulatory reflexes in the assimilating context /pk/ sequences, and we mainly classify lip aperture data from a larger group of subjects.

For the purpose of our current study, we take into account two vocalic contexts (high front vowel /i/-to-/i/ vs. low central vowel /a/-to-/a/), two speech rates (e.g., normal vs. fast), and assimilating /pk/ sequences along with two homorganic controls (/pp/ and /kk/). Note that American English is relatively well studied in terms of numerous vocalic contexts (cf., an aerodynamic study on Korean place-assimilating sequences using various vocalic contexts in Jun (1995)). Various vowels (e.g., non-high back vowels as in 'pop', 'tot', and 'caulk') were used in Browman & Goldstein (1995) and a high front lax vowel (or a schwa) (e.g., 'perfect') in Browman & Goldstein (1990c). Using gestural simulations, a low front vowel (e.g., 'bad') was used in Byrd (1992) and Chen (2003). In Öhman's (1967) X-ray data, different articulators involved distinct vocalic and consonantal tiers: the tongue tip articulation for voiced stop /d/ fairly consistently achieved its constriction degree regardless of vocalic context (/i/-to-/i/, /a/-to-/a/, /u/-to-/u/) are, and the articulations of the tongue body and the lips for flanking vowels were vowel-dependent. By hypothesis, gestural tiers are divided into the consonantal tier and the vocalic tier (Browman & Goldstein, 1992; Öhman, 1967), where a consonant is superimposed over the vocalic tier. However, invariant target achievement has not been kinematically examined in an assimilating context, where articulatory reduction can potentially be applied. It is of interest to learn whether this invariance which is independent of vocalic context (high vowel /i/-to-/i/ and low vowel /a/-to-/a/) holds true for target achievement in the assimilating /pk/ sequences.

We are also interested in describing vertical jaw movement. The jaw articulator is shared by both vocalic and consonantal tiers (e.g., jaw height for front vowels gradually decreasing in the order /i/ > /ɪ/ > /e/ > /æ/ and for back vowels /u/ > /ʊ/ > /ɑ/ in Ladefoged (2001); jaw height for consonants in the order coronal > (labial, velar) in Keating et al. (1994)). As we examine vocalic context effects (/i/-to-/i/ vs. /a/-to-/a/) on jaw height, we also examine whether jaw height of C1 varies as a function of C2 (e.g., assimilating heterorganic /pk/ sequences vs. control homorganic /pp/ sequences) and speech rates (normal vs. fast). Previous kinematic studies on flapping in a high vowel context ([iri]) (Son, 2015b) and a low-vowel context ([ara]) (Son, 2015a) exhibited no speech rate effects on jaw height. Although previous studies have reported inter-/within-speaker variability with respect

to articulatory reduction of C1 in assimilating /pk/ sequences in terms of primary articulator (e.g., either categorical, gradient, or both) (Jun, 1995; Son, 2008; Son et al., 2007), they are lacking in jaw movement data. In the current study, we examine whether, and if so how, jaw movement varies as a function of different vowel (/i/-to-/i/ vs. /a/-to-/a/), C1C2 type (/pk/ vs. /pp/ vs. /kk/), and speech rate (normal vs. fast) conditions. In doing this, we attempt to provide a more detailed description of C1C2 sequences.

2. Methods

2.1. Data collection and subject

We used an electromagnetic midsagittal articulometer (EMMA in Perkell et al., 1992; for a detailed description of this system and subsequent post-processing procedures, see Son (2008)). We made use of kinematic data from four transducers (the upper and lower lips for lip aperture, the tongue dorsum, and the lower central incisor for the jaw) for further analysis.

Eight native speakers (five females and three males) of Seoul Korean participated in the EMMA experiment. Kinematic data from Son's (2013) /pk/ sequences in the low-vowel (/a/-to-/a/) context at two speech rates (normal vs. fast) were reused. In addition, we examined three consonantal sequences (/pk/ and controls (/pp/ and /kk/)) in the high-vowel (/i/-to-/i/) context as well as control /pp/ and /kk/ sequences in the low-vowel (/a/-to-/a/) context, which were elicited simultaneously at the time of data collection in Son (2013). The subjects, ranging from their mid-twenties to early thirties, were living in Connecticut, U.S.A. when we collected their articulatory data, where they had been pursuing graduate or post-doctoral research. In the pre-experimental, paper-and-pencil questionnaire, all of them identified themselves as native speakers of the Seoul-Korean dialect, reporting that they had lived in the Seoul metropolitan area for at least twenty-three years.¹ They had all spent for approximately four years abroad on average and belonged to Korean communities of various kinds while abroad (e.g., Korean churches, Korean student associations, etc.). None of them had speech/hearing deficits. They were not informed of the purpose of the experiment before or after the experiment and were all financially rewarded after completing a production experiment.

The stimuli list is provided in (1). A total of 256 tokens (2 (Vowel contexts) × 2 (Speech rates) × 8 (Repetitions) × 8 (Subjects)) were available both for /pk/ and /pp/ sequences. A total of 251 tokens were available for /kk/ since we failed to acquire four /kk/ tokens at normal rate and we excluded one /kk/ token at normal rate due to poor trajectory of the tongue dorsum.

1 One female speaker (Subject 1) lived in France between the ages 4 and 7, and at age 11.

(1) Stimuli list

i. Test sequences: /pk/

- a. /ʃən~~ap~~kanin kjesanhaki himtɨlɔ/ (borrowed from Son (2013: 670))
전압가는 계산하기 힘들어.
‘Voltage values are difficult to estimate.’
- b. /ʃən~~ip~~kiɐ ta nawaissɔ/
전입기에 다 나와 있어.
‘Everything (you need) is on the resident document.’

ii. Control sequences: /kk/

- a. /ʃən~~ak~~kanin seloun tehaklo ilimija/
전학가는 새로운 대학로 이름이야.
‘Jeonhakka is a new name for a university road.’
- b. /iik~~kin~~in sɔlapane issɔ/
이익기는 서랍안에 있어.
‘(The) profit-record book is in the drawer.’

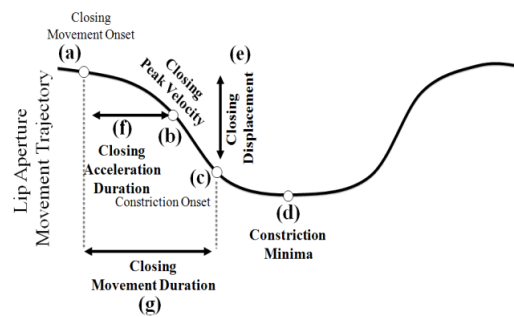
iii. Control sequences: /pp/

- a. /ʃən~~ap~~panin mantin tanɔja/
전압바는 만든 단어야.
‘Jeonappa is a coined word.’
- b. /ʃən~~ip~~pilika simhakun/
전입비리가 심하군.
‘(A) camouflage transfer is too bad.’

2.2. Measurements

We semi-automatically demarcated three gestural landmarks (e.g., the movement onset, peak velocity, and constriction onset) for the lip aperture gesture as we use the function *lp_Findgest* and constriction minima as we use the function *lp_Snapex* in *MVIEW* (Tiede, 2005) (Figure 1.i) (see Son (2011) for more detailed descriptions of data analysis using Tiede’s (2005) algorithm based on a velocity threshold). Using the function *lp_Snapex*, constriction maxima were also demarcated in the tongue dorsum and the jaw gestures. Constriction minima and constriction maxima shown with a red dot are pertinent to maximum constriction in the kinematic trajectories (e.g., the vertical tongue dorsum, lip aperture, and vertical jaw position) (Figure 1.ii). When the function *lp_Snapex* failed to capture minimum lip aperture points in time, we used values corresponding to the time point of the maximum constriction of the vertical tongue dorsum trajectory (Figure 1.iii). Based on the measures as shown in figures (1.i), (1.ii), and (1.iii), we borrowed criteria for token classification established in Son (2008)².

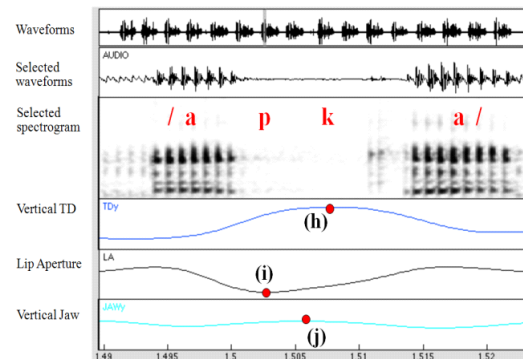
i. Schematized lip aperture movement trajectory marked with independent kinematic variables



Lip aperture minima

- a. Movement onset of the lip aperture closing movement (Fig. 1.i.a).
- b. Peak velocity of the lip aperture closing movement (Fig. 1.i.b).
- c. Constriction onset (i.e., target attainment) of the lip aperture gesture (Fig. 1.i.c).
- d. Constriction minima of the lip aperture gesture (Fig. 1.i.d).
- e. Closing displacement (i.e., spatial difference between the lip aperture closing movement onset and the lip aperture constriction onset) (Fig. 1.i.e).
- f. Closing acceleration duration (i.e., time-to-peak velocity) measured from the lip aperture movement onset to the peak velocity (Fig. 1.i.f).
- g. Closing movement duration measured from the lip aperture movement onset to the constriction onset (Fig. 1.i.g).

ii. Articulatory movement trajectory captured from the window of unreduced /p/ in /canapka/ (Token #3 at normal rate by Subject 6)



iii. Articulatory movement trajectory captured from the window of categorically reduced /p/ in /canapka/ (Token #4 at fast rate by Subject 6)

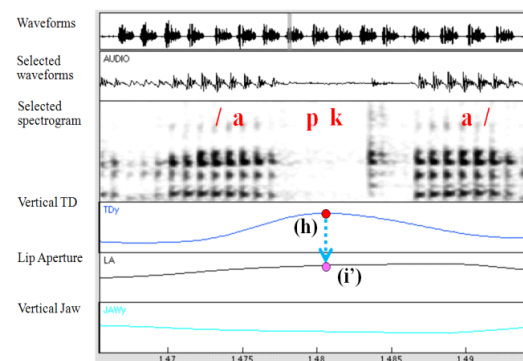


Figure 1. Gestural landmarks measured in CIC2 sequences

² When a lip aperture gesture was not available, a time point of the lip aperture corresponding to the maximum constriction of the vertical tongue dorsum trajectory was used in the current study (Figure 1.iii), instead of the constriction onset in Son (2008: 56-57).

(2) Token classification borrowed after a slight modification from Son (2008: 56-57)

- i. C1 of heterorganic C_iC_k cluster sequences is *unreduced* if:
 - a. *lp_Findgest* detects gestural landmarks (i.e., gesture onset, target attainment, and release) within a properly selected window, and
 - b. Spatial values for maximum constriction are not less constricted than 3 times the standard deviation of the interquartile mean of heterorganic C_iC_k cluster sequences.
- ii. C1 of heterorganic C_iC_k cluster sequences are *partially (or gradually) reduced* if:
 - a. *lp_Findgest* detects gestural landmarks (i.e., gesture onset, target attainment, and release) within a properly selected window, and
 - b. Spatial values for maximum constriction are less constricted than 3 times the standard deviation of the interquartile mean of heterorganic C_iC_k cluster sequences, but more constricted than 3 times the standard deviation of the interquartile mean in homorganic control utterances without the relevant constriction (in C_kC_k control utterances, a time point of the relevant constriction of C1 in C_iC_k was measured at the constriction maxima of homorganic C_kC_k clusters).
- iii. C1 of heterorganic C_iC_k cluster sequences are *categorically reduced* if:
 - a. *lp_Findgest* fails to detect gestural landmarks (i.e., gesture onset, target attainment, and release) within a properly selected window, or
 - b. Spatial values for maximum constriction are less constricted than 3 times the standard deviation of the interquartile mean of heterorganic C_iC_k cluster sequences, and not more constricted than 3 times the standard deviation of the interquartile mean in homorganic control utterances without the relevant constriction (in C_kC_k control utterances, a time point of the relevant constriction of C1 in C_iC_k was measured at the constriction maxima of homorganic C_kC_k clusters).

We also applied the criteria for token classification in (2) to categorize control /pp/ sequences into two gestural types (e.g., unreduced and partially reduced). Doing this for control /kk/ sequences, we categorize tokens as partially reduced if spatial values for the tongue dorsum constriction are less constricted than 3 times the standard deviation of the interquartile mean.

2.3. Statistical analysis

Raw data were converted to z-scores and used as input for further analysis. Linear mixed-effects models were constructed in R (R Development Core Team, 2014). The results of articulatory analysis were fitted with the *lmer* function from the *lme4* packages (Bates et al., 2015). Specifically, we fitted a linear regression model with four kinematic measurements of the lip aperture

closing movement (peak velocity, spatial displacement, acceleration duration, and closing movement duration), minimum lip aperture, and maximum vertical jaw position while taking into account Vowel context (low vowel /a/-to-/a/ vs. high vowel /i/-to-/i/), Speech rate (normal vs. fast), and Consonant sequence type (control /kk/ vs. assimilating /pk/ vs. control /pp/). Tukey HSD tests were used for post-hoc analysis.

3. Results³

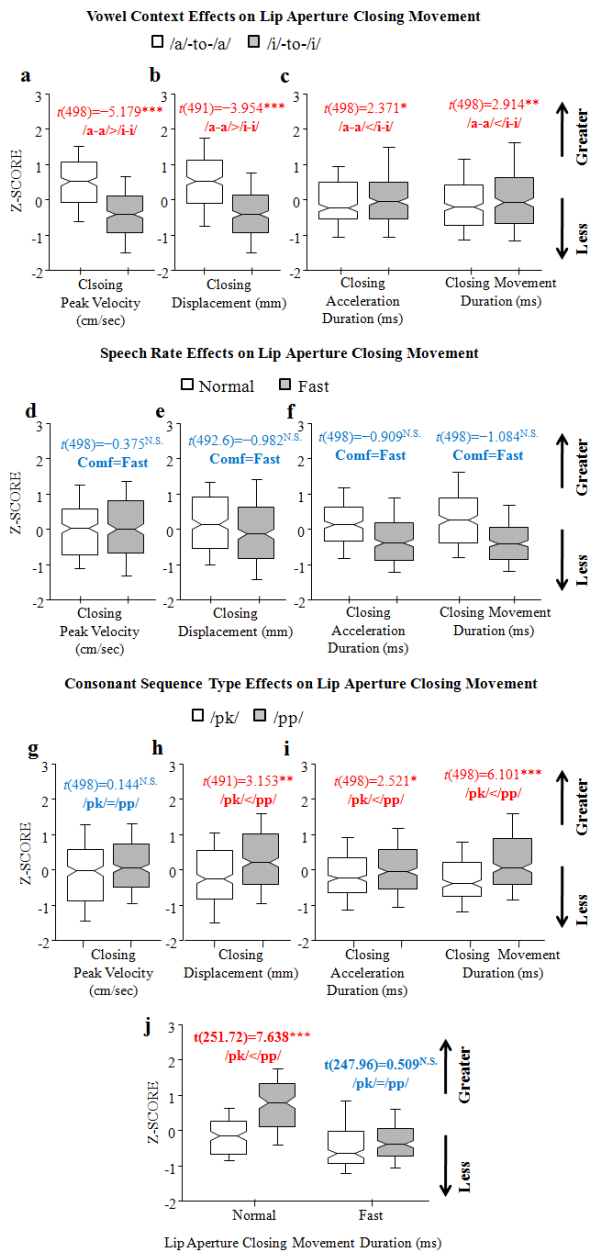
3.1. Lip aperture closing movement of /pk/ sequences in comparison with /pp/ sequences

With regard to a main effect of Vowel context (/i/-to-/i/ vs. /a/-to-/a/), the peak velocity and spatial displacement were smaller in the high-vowel context, with a reduction of -0.797 ($SE \pm 0.154$) and of -0.613 ($SE \pm 0.155$), respectively [$t(498) = -5.179$; $t(491.1) = -3.954$, all at $p < 0.0001$] (/a/-to-/a/ > /i/-to-/i/) (Figures 2.a and 2.b). Meanwhile, the duration of the lip aperture closing acceleration and the lip aperture closing movement were longer in the high-vowel context, lengthened by 0.397 ($SE \pm 0.167$) and by 0.453 ($SE \pm 0.156$), respectively [$t(498) = 2.371$, $p < 0.05$; $t(498) = 2.914$, $p < 0.01$] (/a/-to-/a/ < /i/-to-/i/) (Figure 2.c). As for a main effect of Speech rate, there was no significant effect on any dependent variables (all at $p > 0.05$) (Figures 2.d, 2.e., and 2.f).

The results of lip aperture closing movement also showed a main effect of Consonant sequence type on spatial displacement, acceleration duration, and movement duration [$t(491.1) = 3.153$, $p < 0.01$; $t(498) = 2.521$, $p < 0.05$; $t(498) = 6.101$, $p < 0.0001$], where /pp/ sequences exhibited greater vertical displacement by 0.489 ($SE \pm 0.155$) and greater duration by 0.422 ($SE \pm 0.167$) and 0.949 ($SE \pm 0.156$), respectively (/pk/ < /pp/) (Figures 2.h and 2.i). There was an interaction between Consonant sequence type and Speech rate on the duration of the lip aperture closing movement, with a reduction of -0.642 ($SE \pm 0.223$) in /pp/ sequences when combined with fast rate [$t(498) = -2.881$, $p < 0.01$], where shorter duration in /pk/ sequences was observed only in the normal rate (/pk/ < /pp/), and not in the fast rate (/pk/ = /pp/) (Figure 2.j).

Taking everything into account, the peak velocity only varied with vocalic environments. Spatiotemporal properties showed the opposite patterns of one other with respect to vocalic contexts: greater spatial displacement was observed in the low-vowel (/a/-to-/a/) context and temporally longer closing duration in the high-vowel (/i/-to-/i/) context. Lastly, spatiotemporal reduction during the lip aperture closing movement was generally observed in the assimilating /pk/ sequences in comparison with control /pp/ sequences (/pk/ < /pp/) (Figure 2.i), with one exception (e.g., the duration of lip aperture closing movement in fast rate (/pk/ = /pp/) (Figure 2.j).

³ The results of linear mixed-effects models are shown in the appendix.



* $p < .05$, ** $p < .01$, *** $p < .0001$.

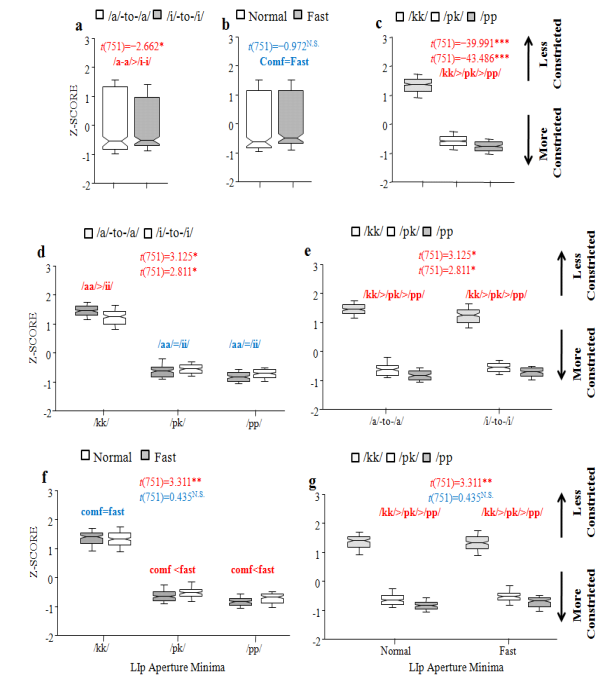
Figure 2. Lip aperture measures during the closing movement (the peak velocity, spatial displacement, acceleration duration, and movement duration) as a function of (a, b, c) Vowel context [the top panel], (d, e, f) Speech rate [the second panel], (g, h, i) Consonant sequence type [the third panel], and (j) Interaction between Speech rate and Consonant sequence type [the bottom panel].

3.2. Gestural reduction of lip aperture

3.2.1. Lip aperture minima of /pk/ sequences in comparison with /pp/ and /kk/ controls

Regarding lip aperture minima, there was interaction between Vowel context and Consonant sequence type, with a reduction of 0.236 ($SE \pm 0.075$) in /ipki/ and of 0.212 ($SE \pm 0.075$) in /ippi/ sequences [$t(751) = 3.125$; $t(751) = 2.811$, all at $p < 0.01$], and between Speech rate and Consonant sequence type, with a reduction of 0.248 ($SE \pm 0.075$) in /pk/ sequences when combined with fast rate [$t(751) =$

3.311, $p < 0.001$]. The maximum constriction degree between the upper and lower lips did not vary with different vocalic contexts within the assimilating context (/apka/=/ipki/) (Figure 3.d). This was also true for homorganic control /pp/ sequences (/appa/=/ippi/). In addition, there were Speech rate effects on the lip aperture minima as long as /p/ was included (normal (more constriction) < fast (less constriction) in both /VpkV/ and /VppV/), indicating that the fast rate is associated with less constriction (Figure 3.f). With regard to interaction between the two factors, from another perspective, the lip aperture was consistently less constricted in the order /kk/ (less constriction) > /pk/ > /pp/ (more constriction) in each combination (Figures 3.e (interacting with Vowel context) & 3.g (interacting with Speech rate)).



* $p < .01$, ** $p < .001$, *** $p < .0001$.

Figure 3. Lip aperture minima as a function of (a) Vowel context [the top panel at left], (b) Speech rate [the top panel at center], and (c) Consonant sequence type [the top panel at right]. Interaction between Vowel context and Consonant sequence type is plotted by (d) consonantal sequence types and (e) vocalic contexts. Interaction between Speech rate and Consonant sequence type is plotted by (f) consonantal sequence types and (g) speech rates.

3.2.2. Reduction frequency in each cluster type

Referring to the criteria established to classify partially (or gradually) reduced tokens (see (2) in section 2.2), we learned that 49 tokens were partially reduced and 6 tokens categorically reduced out of a total of 763 tokens. As shown in Table 1, partial (or gradient) reduction occurs dominantly in fast rate in 82% of tokens (40 out of 49) and categorical reduction in 100% of tokens (6 out of 6). With respect to vocalic environment, the high-vowel context is a condition for more frequent partial reduction in 59% of tokens (29 out of 49), compared to the low vowel in 41% (20 out of 49). As for Consonant sequence type, not a single sequence type demonstrated a dominant reduction frequency over the others in terms of partial reduction (37% (18 out of 49 tokens) for /kk/ sequences; 33% (16 out of 49

tokens) for /pk/ sequences; 31% (15 out of 49 tokens) for /pp/ sequences). Categorical reduction was typical of the assimilating /pk/ sequences. If homorganic control sequences showed any instance of gestural reduction, so did the assimilating /pk/ sequences.

Table 1. Reduction frequency based on the criteria in (2).

Note that ‘N’ is an abbreviation for normal rate and ‘F’, fast rate. The instances of partial (or gradient) reduction are shown in red and so does categorical reduction in parentheses. On the bottom row in isolation, the probability is also shown in angle brackets

CC	/kk/ <251 tokens>				/pk/ <256 tokens>				/pp/ <256 tokens>			
	/i/-to-/i/		/a/-to-/a/		/i/-to-/i/		/a/-to-/a/		/i/-to-/i/		/a/-to-/a/	
sr	N	F	N	F	N	F	N	F	N	F	N	F
K1	2	0	0	2	0	2	0	2	0	1	0	0
K2	0	2	1	0	0	2	1	1	0	0	0	0
K3	0	0	0	2	0	0	0	(1)	1	5	0	0
K4	0	0	0	0	0	1	0	1(4)	2	1	0	0
K5	0	0	0	1	0	2	0	2	0	2	0	0
K6	1	1	0	3	0	0	0	(1)	0	0	0	0
K7	0	0	0	0	0	0	0	0	0	0	0	0
K8	0	0	0	3	0	2	0	0	1	1	0	1
Sum	3	3	1	11	0	9	1	6(6)	4	10	0	1
fast	4				1				4			
Nrml	14				15(6)				11			
/i/-/i/	6				9				14			
/a/-/a/	12				7(6)				1			
Total	18 <7.2%>				16 <6.3%> (6 <2.3%>)				15 <5.9%>			

3.3. Vertical jaw position

With regard to vertical jaw position, there was a main effect of Vowel context, with an increase of 0.967 ($SE \pm 0.162$) in the high-vowel context [$t(647) = 5.984, p < 0.0001$] (/a/-to-/a/ </i/-to-/i/) (Figure 4.a), Speech rate, with a decrease of -0.424 ($SE \pm 0.163$) in the fast fate [$t(647) = -2.609, p < 0.01$] (normal > fast) (Figure 4.b), and Consonant sequence type, with an increase of 0.611 ($SE \pm 0.148$) in the /pk/ sequence and of 0.358 ($SE \pm 0.147$) in the /pp/ sequence [$t(647) = 4.124, p < 0.0001$; $t(647) = 2.441, p < 0.05$] (/kk/ </pk/ = /pp/) (Figure 4.c). Specifically, vertical jaw position was higher when a C1C2 sequence was flanked by high vowels (/i/-to-/i/). As for different consonantal sequence types, they demonstrated binary distribution (/kk/ </pk/ = /pp/) (cf., ternary distribution for maximum lip constriction (/kk/ </pk/ </pp/)). Lastly, vertical jaw position varied with speech rate, indicating that the mouth is more open when speakers talk faster.

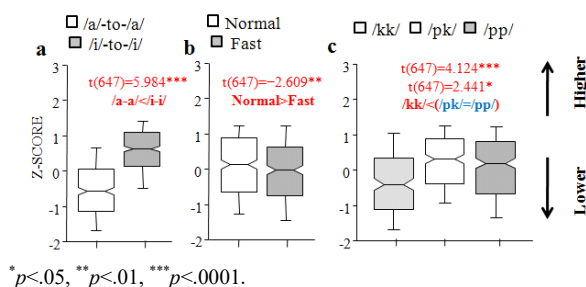


Figure 4. Vertical jaw position as a function of (a) Vowel context, (b) Speech rate, and (c) Consonant sequence type.

4. Discussion

4.1. Spatial reduction in magnitude as a component of the quantitative properties of articulatory gestures, unlikely to be an independent reason for place assimilation

Examining lip aperture maxima, results from the current study indicated that the number of tokens classified as partially reduced were very similar across different consonantal sequence types (18 tokens for control /kk/ sequences vs. 16 tokens for test /pk/ sequences vs. 15 tokens for control /pp/ sequences). Although we were not able to separate the first consonant from the second in control $C_i C_i$ (/pp/) and $C_k C_k$ (/kk/) sequences in terms of maximum constriction, the maximum constriction of these homorganic sequences is at least not considered to be reduced but fortified, compared to lenis (/pk/). Evaluated within the same cluster sequence type, the probability across different sequences was quite comparable (6.3% for the test /pk/ sequences vs. 5.9% and 7.2% for the control /pp/ and /kk/ sequences, respectively). Based on this, we tentatively conclude that i) partial reduction occurs to some extent across the board, if there is any, and ii) gradient reduction in magnitude may be attributed to the general quantitative properties of gestures (Browman & Goldstein, 1986, 1989), rather than being specific to place-assimilating sequences such as /pk/ (e.g., Jun, 1996).

However, caution should be taken before we come to a conclusion since this is a preliminary study which covers only a subset of data: the current analysis only includes the within-word boundary condition, and not the across-word boundary condition. Another disadvantage may be stem from following Son's (2008) arbitrary requirements for categorizing partially (or gradiently) reduced tokens, which referred to an interquartile mean and interquartile standard deviation in search of their resistance to outliers. Use of an interquartile mean and three times its standard deviation to define partial (or gradient) reduction still remains an arbitrary method. An alternative can be found in Son et al. (2012), where frequency distributions are interpreted using a histogram as they provide a mathematical analysis based on point-attractors in the task-dynamics model of speech production (Nam et al., 2012).

4.2. Spatiotemporal weakening of closing movement in the assimilating /pk/ sequences

In the current study, several kinematic properties of the lip aperture gesture reflected weaker articulation in the spatiotemporal dimension (e.g., closing spatial displacement, closing acceleration duration, and closing movement duration) in the assimilating context (/pk/ </pp/). Traditionally, Korean place assimilation is driven by applying a feature-changing rule as shown in (3). Using narrow phonetic transcription, two phonological processes are involved in Korean assimilating /pk/ sequences. One is a fortification rule in which a lenis obstruent in the onset is fortified after an obstruent in the coda (Silverman, 2017). The other is a place-assimilation rule in which the place feature of the coda becomes identical with that of the onset (Kim-Renaud, 1974). In addition, fortification occurs regardless of the application of the place-assimilating rule and the lenis stop is assumed to be realized in the coda (e.g., /pk/ → [kk*]; /pp/ → [pp*]).

(3) Feature-changing rule application

Input /tapkəl/ ‘reply’

Fortification [-asp, -tense]→[+tense] / [-asp, -tense] ___
Feature change [-cont, +ant]→[i Place] / ___ [i Place]

Output [takəl]

- The symbol [i] stands for a shared feature.

Recall that spatiotemporal reduction is not peculiar to the heterorganic assimilating context but is even applied to homorganic controls. We further fitted a linear regression model for the three dependent variables after excluding reduced tokens of any kind. The results showed that a weaker gestural event of C1 from heterorganic assimilating sequences was also consistently observed [$t(460.4)=3.096$, $p<0.01$ for the closing displacement; $t(467)=2.511$, $p<0.05$ for the closing acceleration duration; $t(467)=6.124$, $p<0.0001$ for the closing movement duration] (/pk/ < /pp/). Given this, we speculate on two possibilities to account for this. One is that spatiotemporal reduction observed even in *unreduced* tokens generally occurs at the phonetic execution level. Speakers’ preemptive action with more reduction in C1 may, in part if not completely, reflect speakers’ strategy that they simply do not exert much articulatory efforts towards a perceptually unrewarding gestural event. At the phonological level of representation, this reduction could be a consequence associated with gestural overlap or independent of gestural overlap in the gestural score (Chen, 2003; Browman & Goldstein, 1990a, 1990b, 1990c; among others). Alternatively, this could be an output of the application of probability-based constraints in optimality theory (Jun, 1995). Since it is beyond the scope of the current study to discuss which theoretical frameworks are a better fit to account for data we have acquired, we leave this issue for further study. The other possibility to account for a weaker gestural event of C1 being consistently observed from heterorganic assimilating sequences is that stronger articulation of fortified onset [p*] may have extended onto coda [p] in the homorganic C1C2 sequences. Note that Barry (1991) observed in his electropalatography study that in an assimilating sequence from English there was a 43% increase in fast rate, compared to slow rate, in terms of the duration of the rear tongue body gesture in C2 and a 78% increase in terms of the duration of velar closure in C2. (e.g., ‘*hand grenade*’). Given that, it is plausible to assume that C2 in /pp/ sequences may enhance articulatory strength not only its temporal domain but also spatial domain over C1 by extending its articulatory strengthening to its homorganic C1 (e.g., gestural blending in Romero (1992)). As such, strengthening may be spatiotemporally overriding the closing movement of C1.

Caution should be taken, however, since similar spatiotemporal properties in the lip aperture closing movement are observed between lenis and fortis when a single segment is tested in /a/-to-/a/ and /i/-to-/i/ contexts (/VpV/ = /Vp*V/ in Son et al. (2012)). We should include a more balanced set of articulatory data so that we may test the nature of coda in intervocalic homorganic lenis-lenis (e.g., assimilating and non-assimilating) and lenis-fortis (e.g., homorganic and heterorganic) sequences, as well as

in an intervocalic singleton fortis stop control. Doing this will enable a better understanding of articulatory characteristics of lenis in coda, irrespective of the target of place assimilation.

4.3. Gestural reduction in the target of place assimilation attributed to a more open oral tract

Öhman (1967) observed that consonantal articulation (e.g., coronal) is achieved, regardless of different vocalic contexts. Being compatible with this, Korean /pk/ sequences also demonstrated invariable maximum constriction degrees between two different vocalic contexts (/apka/ = /ipki/). This was also true for homorganic control /pp/ sequences (/appa/ = /ippi/). The results of the current study support that articulatory tiers are bifurcated so that consonantal articulation occupies a separate tier, independent of vocalic articulation (Browman & Goldstein, 1992; Öhman, 1967). In particular, consonantal articulation is less open than vocalic articulation in terms of constriction degree in the oral tract. Focusing within a single tier of articulation (e.g., the vocalic tier), low vowels, for example, are less constricted than high vowels (Ladefoged, 2001).

Although the scope of the current study is limited morphologically to the within-word condition, categorical reduction of the lip aperture gesture did not at all occur in the high-vowel context (/ipki/), though we sometimes observed such cases in the low-vowel context (/apka/) in conjunction with fast rate. This elision indicates that speakers are sensitive not only to paralinguistic factors such as speech rates (see speech rate/style formally incorporated within a probability-based, optimality-theoretic approach in Jun (2004); frequency effects in Pouplier et al. (2017)), but also linguistic factors such as vocalic contexts. With respect to occasional categorical reduction, one can conjecture that speakers are aware that it may require too much effort to complete a lip aperture gesture from lower jaw position (e.g., /a/-to-/a/), especially in conjunction with fast rate. At the speech-planning level, therefore, segmental deletion happens as a consequence of this (cf., residual tongue tip gesture detected in C1 for American English in Browman & Goldstein (1990c)). Under this assumption, articulatory efforts can be maximally reserved when the energy cost of reaching a target is unrewarding from the standpoint of speakers, although inter-/intra-speaker variability still remains. Referring to the results of the current study, a significant increase in energy may signal a greater peak velocity and greater spatial displacement in the low-vowel context (e.g., /a/-to-/a/ > /i/-to-/i/).

The occurrence of partial (or gradient) as well as categorical reduction in the lip aperture gesture was obviously rate-dependent, demonstrating higher frequency in fast rate (82% for partial reduction and 100% for categorical reduction as shown in Table 1). This is compatible with a general observation where lenition is more likely to occur easily in fast rate and the vocal tract is more open in this condition (e.g., *close* (unfricated) approximant in slow natural speech vs. *more open* approximant in *faster* speech rate in Kirchner (1998: 257)). Not having acoustic descriptions available which correspond to all articulatory data used in the current study, we are not ready to provide a fully balanced acoustic analysis of C1 /p/ along with articulatory counterparts. Nevertheless, in conformity with Kirchner’s (1998) observation of more *open* articulation in *faster* speech rate, we found that the jaw is also lower in fast rate (normal > fast), in

addition to overall rate effects on the lip aperture minima in the assimilating /pk/ and control /pp/ sequences. This indicates that speech-rate effects are also immediately reflected in jaw articulation as well as the primary articulator. Through this, the spatially reduced mandibular cycle (see *basic mandibular cycle* for consonants, vowels, and syllables, etc. (MacNeilage & Davis (1993: 341)) can facilitate the production of a syllable (e.g., V-to-C) in fast rate, which contributes to a reserve of articulatory efforts as a consequence (Jun, 1996).

Acknowledgments

I wish to thank Louis Goldstein for his invaluable suggestions and corrections in all stages of research on Korean place assimilation, which is rooted in his original proposal for Articulatory Phonology. I am also grateful to three anonymous *Phonetics and Speech Sciences* referees for their detailed comments, eight study subjects for their participation in the EMMA experiments, and Sean C. O'Rourke for his proofreading of this paper. This preliminary study is based on a larger-scale project with Jongho Jun and Hosung Nam on Korean place assimilation supported by NIH Grant DC-00403 to Catherine T. Best (PI) and Haskins Laboratories. Any remaining errors are my own.

References

- Barry, M. C. (1991, August). Temporal modelling of gestures in articulatory assimilation. *Proceedings of the 12th International Congress of Phonetic Sciences* (Vol. 4, pp. 14-17). Aix en Provence, France.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Borden, G. J., & Harris, K. S. (1984). *Speech science primer: Physiology, acoustics, and perception of speech* (2nd ed.). Baltimore, MD: Williams & Wilkins.
- Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.
- Browman, C., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6(2), 201-251.
- Browman, C., & Goldstein, L. (1990a). Gestural structures: Distinctiveness, phonological processes, and historical change. Modularity and the motor theory of speech perception. *Proceedings of a Conference to Honor Alvin M. Liberman* (pp. 313-338). Lawrence Erlbaum Associates, Mahwah, NJ.
- Browman, C., & Goldstein, L. (1990b). Representation and reality: physical systems and phonological structure. *Journal of Phonetics*, 18(3), 411-424.
- Browman, C. P., & Goldstein, L. (1990c). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 341-376). Cambridge, UK: Cambridge University Press.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155-180.
- Browman, C. P., & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti, & L. J. Raphael (eds.), *Producing speech: Contemporary issues (for K.S. Harris)* (pp. 19-33). Woodbury, NY: AIP Press.
- Browman, C. P., & Goldstein, L. Articulatory phonology. Unpublished manuscript.
- Byrd, D. (1992). Perception of assimilation in consonant clusters: a gestural model. *Phonetica*, 49(1), 1-24.
- Chen, L. H. (2003). The origins in overlap of place assimilation. *Proceedings of the XXIIth West Coast Conference on Formal Linguistics* (pp. 137-150). Cascadilla Press, Somerville, MA.
- Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2, 225-252.
- DePaul, R., & Abbs, J. H. (1996). Quantitative morphology and histochemistry of intrinsic lingual muscle fibers in Macaca fascicularis. *Acta Anat*, 155(1), 29-40.
- Fujimura, O., Macchi, M. J., & Streeter, L. A. (1978). Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech*, 21(4), 337-346.
- Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach* (Ph.D. dissertation). University of California in Los Angeles.
- Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (eds.), *Phonetically based phonology* (pp. 58-86). Cambridge, UK: Cambridge University Press.
- Keating, P. A., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, 22(4), 407-422.
- Kim-Renaud, Y-K. (1974). *Korean consonantal phonology* (Ph.D. dissertation). University of Hawaii.
- Kirchner, R. M. (1998). *An effort-based approach to consonant lenition* (Ph.D. dissertation). University of California in Los Angeles.
- Kochetov, A., & Pouplier, M. (2008). Phonetic variability and grammatical knowledge: an articulatory study of Korean place assimilation. *Phonology*, 25(3), 399-431.
- Krakow, R. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures* (Ph.D. dissertation). Yale University.
- Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell.
- MacNeilage, P. F., & Davis, B. L. (1993). Motor explanation of babbling and early speech patterns. In: B. De Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 341-352). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Öhman, S. E. G. (1967). Numerical model of coarticulation. *Journal of Acoustical Society of America*, 41, 310-320.
- Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., & Jackson, M. T. (1992). Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America*, 92, 3078-3096.
- Pouplier, M., Marin, S., Hoole, P., & Kochetov, A. (2017). Speech rate effects in Russian onset clusters are modulated by frequency, but not auditory cue robustness. *Journal of Phonetics* 64, 108-126.
- R Development Core Team. (2014). R: A language and environment for statistical computing. [Computer software]. Vienna,

- Austria: R Foundation for Statistical Computing. Retrieved from <http://www.rproject.org/>
- Romero, J. (1992). An experimental analysis of spirantization in Spanish. *The Journal of the Acoustical Society of America*, 92, 2340.
- Silverman, D. (2017). *A critical introduction to phonology: Functional and usage-based perspectives* (2nd ed.). London, UK: Bloomsbury Academic.
- Son, M. (2008). Gradient reduction of C₁ in /pk/ sequences. *Phonetic Sciences*, 15(4), 43-65.
- Son, M. (2011). The projection of syllable structure: A case study of intervocalic /k/ in Korean. *The Journal of Linguistics*, 36(2), 395-414.
- Son, M. (2013). Kinematics of Korean labial stop /p/ in assimilating and nonassimilating contexts. *The Journal of Studies in Language*, 28(4), 743-765.
- Son, M. (2015a). Articulatory properties of the allophonic variant [ɾ] in Korean /l/-flapping: Gestural reduction and the role of gestural overlap. *Studies in Phonetics, Phonology, and Morphology*, 21(3), 427-456.
- Son, M. (2015b). Korean /l/-flapping in an /i/-i/ context. *Phonetics and Speech Sciences*, 7(1), 151-163.
- Son, M., Alexei K., & Marianne P. (2007). The role of gestural overlap in perceptual place assimilation in Korean. In J. Cole, & J. I. Hualde (eds.), *Papers in laboratory phonology IX* (pp. 507-534). New York, NY: Mouton de Gruyter.
- Son, M., Kim, S., & Cho, T. (2012). Supralaryngeal articulatory signatures of three-way contrastive labial stops in Korean. *Journal of Phonetics*, 40(1), 92-108.
- Son, M., Nam, H., & Jun, J. (2012). Dynamics of place assimilation: A case study in Korean. Poster presented at *the 13th Conference on Laboratory Phonology*.
- Steriade, D. (2000). *The phonology of perceptibility effects: the P-map and its consequences for constraint organization* (Unpublished manuscript). UCLA, Los Angeles, CA.
- Steriade, D. (2001). Directional asymmetries in place assimilation: a perceptual account. In: V. Hume, & K. Johnson (eds.), *The role of speech perception in phonology* (pp. 219-250). San Diego, CA: Academic Press.
- Tiede, M. (2005). *MVIEW: Software for visualization and analysis of concurrently recorded movement data*. New Haven, CT: Haskins Laboratories.

• **Minjung Son**, Corresponding Author
 Associate Professor, Dept. of English Language and Literature
 Hannam University
 70 Hannam-ro, Daedeok-gu, Daejeon 34430, Korea
 Tel: +82-42-629-7342
 Email: minjungson@hnu.kr
 Areas of interest: Articulatory phonology, Articulatory phonetics

Appendix

Table A6. Results of linear mixed-effects models on vertical jaw position

Table A1. Results of linear mixed-effects models on the peak velocity of the lip aperture

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	0.274	0.111	211.20	2.459	0.01474*
CC [pp]	0.489	0.155	491.10	3.153	0.00171**
Vowel [ii]	-0.613	0.155	491.10	-3.954	<81e-05***
SR [fast]	-0.156	0.159	492.60	-0.982	0.32653
CC[pp]:V[ii]	-0.240	0.219	491.10	-1.095	0.27393
CC[pp]:SR[fast]	0.170	0.222	491.90	0.765	0.44453
V[ii]:SR[fast]	-0.356	0.222	491.90	-1.604	0.10945

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	-0.715	0.108	647.00	-6.646	<6.41e-11***
CC [pk]	0.6114	0.148	647.00	4.124	<4.20e-05***
Vowel [pp]	0.358	0.147	647.00	2.441	0.01493*
Vowel [ii]	0.967	0.162	647.00	5.984	<3.61e-09***
SR [fast]	-0.424	0.163	647.00	-2.609	0.00929**
CC [pk] : V [ii]	-0.208	0.216	647.00	-0.965	0.33508
CC [pp] : SR [ii]	0.149	0.216	647.00	0.689	0.49115
CC [pk] : SR[fast]	0.114	0.218	647.00	0.521	0.60247
CC [pp] : SR[fast]	0.144	0.216	647.00	0.668	0.50424
V [ii] : SR [fast]	0.235	0.233	647.00	1.011	0.31254

Table A2. Results of linear mixed-effects models on the spatial displacement of the lip aperture

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	0.396	0.109	498.00	3.643	0.000298**
CC [pp]	0.022	0.154	498.00	0.144	0.885498
Vowel [ii]	-0.797	0.154	498.00	-5.179	<3.24e-07***
SR [fast]	-0.059	0.158	498.00	-0.375	0.707800
CC [pp] : V [ii]	-0.075	0.218	498.00	-0.344	0.730716
CC [pp] : SR[fast]	0.343	0.220	498.00	1.557	0.120116
V [ii] : SR [fast]	-0.272	0.220	498.00	-1.234	0.217691

Table A3. Results of linear mixed-effects models on the duration of lip aperture closing acceleration

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	-0.161	0.118	498.00	-1.361	0.1741
CC [pp]	0.422	0.167	498.00	2.521	0.0120*
Vowel [ii]	0.397	0.167	498.00	2.371	0.0181*
SR [fast]	-0.156	0.172	498.00	-0.909	0.3640
CC [pp] : V [ii]	-0.249	0.237	498.00	-1.050	0.2941
CC [pp] : SR[fast]	-0.240	0.240	498.00	-1.002	0.3167
V [ii] : SR [fast]	-0.096	0.240	498.00	-0.401	0.6888

Table A4. Results of linear mixed-effects models on the duration of lip aperture closing movement

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	-0.287	0.110	498.00	-2.605	0.00945**
CC [pp]	0.949	0.156	498.00	6.101	<2.12e-09***
Vowel [ii]	0.453	0.156	498.00	2.914	0.00373**
SR [fast]	-0.173	0.160	498.00	-1.084	0.27875
CC [pp] : V [ii]	-0.299	0.220	498.00	-1.360	0.17459
CC [pp] : SR[fast]	-0.642	0.223	498.00	-2.881	0.00414**
V [ii] : SR [fast]	-0.197	0.223	498.00	-0.882	0.37821

Table A5. Results of linear mixed-effects models on lip aperture minima

	Estimate	SE	df	t-value	Pr (> t)
(intercept)	1.441	0.037	751.00	38.525	<2e-16***45
CC [pk]	-2.115	0.053	751.00	-39.991	<2e-16***
Vowel [pp]	-2.299	0.053	751.00	-43.486	<2e-16***
Vowel [ii]	-0.143	0.054	751.00	-2.662	0.007944
SR [fast]	0.051	0.053	751.00	0.972	0.331511
CC [pk] : V [ii]	0.236	0.075	751.00	3.125	0.001844**
CC [pp] : SR [ii]	0.212	0.075	751.00	2.811	0.005061**
CC [pk] : SR[fast]	0.248	0.075	751.00	3.311	0.000974***
CC [pp] : SR[fast]	0.033	0.075	751.00	0.435	0.663629
V [ii] : SR [fast]	-0.147	0.076	751.00	-1.940	0.5125