Coordinative movement of articulators in bilabial stop /p/*

Son, Minjung**

Abstract

Speech articulators are coordinated for the purpose of segmental constriction in terms of a task. In particular, vertical jaw movements repeatedly contribute to consonantal as well as vocalic constriction. The current study explores vertical jaw movements in conjunction with bilabial constriction in bilabial stop /p/ in the context /a/-to-/a/. Revisiting kinematic data of /p/ collected using the electromagenetic midsagittal articulometer (EMMA) method from seven (four female and three male) speakers of Seoul Korean, we examined maximum vertical jaw position, its relative timing with respect to the upper and lower lips, and lip aperture minima. The results of those dependent variables are recapitulated in terms of linguistic (different word boundaries) and paralinguistic (different speech rates) factors as follows. Firstly, maximum jaw height was lower in the across-word boundary condition (across-word < within-word), but it did not differ as a function of different speech rates (comfortable = fast). Secondly, more reduction in the lip aperture (LA) gesture occurred in fast rate, while word-boundary effects were absent. Thirdly, jaw raising was still in progress after the lips' positional extrema were achieved in the within-word condition, while the former was completed before the latter in the across-word condition. Lastly, relative temporal lags between the jaw and the lips (UL and LL) were more synchronous in fast rate, compared to comfortable rate. When these results are considered together, it is possible to posit that speakers are not tolerant of lenition to the extent that it is potentially realized as a labial approximant in either word-boundary condition while jaw height still manifested lower jaw position in the across-word boundary condition. Early termination of vertical jaw maxima before vertical lower lip maxima across-word condition may be partly responsible for the spatial reduction of jaw raising movements. This may come about as a consequence of an excessive number of factors (e.g., upper lip height (UH), lower lip height (LH), jaw angle (JA)) for the representation of a vector with two degrees of freedom (x, y) engaged in a gesture-based task (e.g., lip aperture (LA)). In the task-dynamic application toolkit, the jaw angle parameter can be assigned numerical values for greater weight in the across-word boundary condition, which in turn gives rise to lower jaw position. Speech rate-dependent spatial reduction in lip aperture may be able to be resolved by means of manipulating activation time of an active tract variable in the gestural score level.

Keywords: jaw, jaw angle, lip aperture, lower lip, upper lip, spatial reduction, temporal lag, task-dynamic, activation time

1. Introduction

In articulatory phonology, the basic linguistic unit is a *gesture* that is hypothesized to be abstract, invariant, and physical at the

phonological level of representation (Browman & Goldstein, 1986, 1989, 1992; *inter alia*). To quote Browman & Goldstein (1989), "...gestures are the basic atoms of phonological structures (p.201)." and "gestures are units of action that can be identified by observing the coordinated movements of the vocal tracts (p.202)." In Saltzman & Kelso (1987), a gesture involves articulators which are assembled in a coordinative manner to accomplish a linguistically meaningful vocal tract action. In Browman & Goldstein (1986), *task*-controlled gestures hypothesize to be specified for two *task* variables: constriction location (CL) and constriction degree (CD). The lips are specified as lip protrusion (LP or PRO) and lip aperture (LA), the tongue tip as tongue tip constriction location (TTCL) and tongue tip constriction degree (TTCD), the tongue body as tongue body constriction location (TDCL) and tongue body constriction degree (TDCD), the velum as velic aperture (VEL), and the glottis as glottal aperture (GLO). It is further hypothesized that active tract variables are constructed into a larger coordinative structure, a *gestural score*, and temporal intervals of time are specified for a given target utterance. For a given gesture, a set of articulators engaged in the *task*-specific tract variable are defined in the computational model as shown in Figure 1 (Browman & Goldstein, 1989).

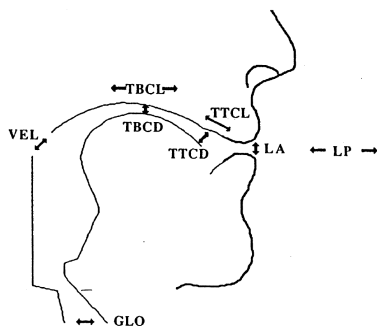| Tract variable	Articulators involved
LP	Upper & Lower lips, Jaw
LA	Upper & Lower lips, Jaw
TTCL	Tongue tip, Tongue body, Jaw
TTCD	Tongue tip, Tongue body, Jaw
TBCL	Tongue body, Jaw
TBCD	Tongue body, Jaw
VEL	Velum
GLO	Glottis

Figure 1. Tract variables and contributory articulators of the computational model borrowed from Browman & Goldstein (1989:207)).

The current study examines bilabial constriction and the jaw. As we limit the scope of investigation to the set of associated articulators for a track variable, lip aperture (LA), we aim to improve general understanding of speech articulation in terms of interarticulator, spatio-temporal coordination for a given speech *task*. We begin by examining lip aperture (LA) and jaw movement engaged in constriction of the lips in the intervocalic labial stop /p/ of Korean so that it can serve as a base for robotic controls in a *task-dynamic*-based model articulator (e.g., estimated values to be entered into dynamic parameters). In addition, we look into temporal relations between the jaw and the lips to enhance our understanding of jaw movement in general.

1.1. Overview of task variables and associated joint variables in a task-dynamic model articulator

In Nam's (ms.) overview of *articulating machines*, *task*-based gestures are articulatory movements that form part of the behaviors of the physical system where motion is predicted as a function of time. Equations of motion of this kind can deal with *task*-based endpoints in the two-dimensional space of Euclidean geometry. From the perspective of robotic movement, kinematic conversion occurs by means of mathematically mapping *task* variables (e.g., $\ddot{x}$, $\dot{x}$, x) to joint variables (e.g., $\ddot{\theta}$, $\dot{\theta}$, θ). In particular, the *task-dynamic* model of speech production provides a set of equations (e.g., partial derivative) using several joint variables (e.g., $\ddot{\theta}$, $\dot{\theta}$, θ) and several *task*-based parameters (e.g., mass, damping, and stiffness) (see Saltzman & Kelso (1987), Nam (manuscript) for a detailed review of *task-dynamics* and relevant equations of motion). Tract variables refer to articulatory behaviors from the perspective of human articulation while joint variables refer to robotic articulation. In terms of joint variables in a *task-dynamic application* toolkit (Nam *et al.*, 2012), the dynamic parameters are engaged in manipulating articulatory weight: greater weight values signal the suppression of articulatory movement and smaller weight values, the augmention of articulatory movement.
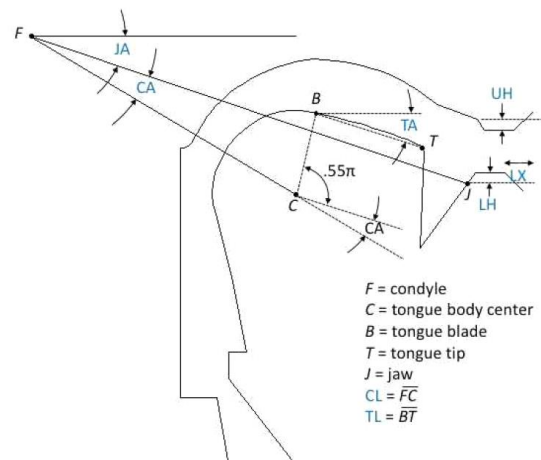

Figure 2. Articulation model (borrowed from Nam (manuscript:2) which is simplified from Mermelstein's original (1973:1071) model-generated vocal tract outline). $\overline{FC}$ indicates the distance from the condyle (F) to the tongue body center (C) and $\overline{BT}$ the distance from the tongue blade (B) to the tongue tip (T)

According to Nam (ms.), Mermelstein's (1973) joint variables are useful for understanding principles of articulatory movement. Mermelstein (1973) proposed a model articulator for which he assigned position variables to fixed and movable compositions of the vocal tract (i.e., the hyoid bone, the jaw, the tongue blade, the tongue body, the lips, the velum, the maxilla, and the pharynx). As shown in Figure 2, the position variables are joint variables expressed in a coordinate plane, which was originally outlined in Mermelstein's model-generated vocal tract; the jaw angle (JA), tongue body center angle (CA), and tongue tip angle (TA) are categorized as revolute joints which provide single-axis rotation movement. The vertical upper lip position (UH), vertical lower lip position (LH), and horizontal lower lip position (LX) are prismatic joints which provide linear sliding movement.

In terms of the mapping relationship between task-controlled tract

variables and joint variables, Nam clarified, in his manuscript, the overview of the *task-dynamics* model (i.e., a model of vocal-tract articulation from the perspective of robotics) where a subset of joint variables is associated with a certain tract variable in articulatory movement. The jaw angle (JA) is repeatedly specified for several tract variables such as lip protrusion (LP or PRO), lip aperture (LA), tongue tip constriction location and degree (TTCL & TTCD), and tongue body constriction location and degree (TDCL & TDCD). In Table 1, each *task*-based tract variable is associated with joint variables.

Table 1. Task-controlled tract variable mapped to associated joint variables (Nam, manuscript: 3)

		Joint Variables									
		LX	UH	LH	JA	CL	CA	TL	TA	NA	GW
Tract Variables	PRO	✓									
	LA		✓	✓	✓						
	TBCL				✓	✓	✓				
	TBCD				✓	✓	✓				
	TTCL				✓	✓	✓	✓	✓		
	TTCD				✓	✓	✓	✓	✓		
	VEL									✓	
	GLO										✓

A task-dynamic application toolkit (TADA) is software *implementing interarticulator speech coordination*, a *coupled oscillator model of intergestural planning*, and a *gestural-coupling model* (Nam *et al.*, 2012). By hypothesis, Browman & Goldstein (1990) proposed that articulatory data from kinematic studies (e.g., using electromagnetic midsagittal articulometer) have been fed dynamic parameter values entered in gestural scores (i.e., constellations of relevant active gestures in the form of a syllable as a basic prosodic unit). Users type in dynamic parameter values for tract variables as well as dynamic parameter values for joint variables to reconstruct area function dynamics and generate acoustic output (Nam, ms.; Nam & Saltzman, 2003; Nam *et al.*, 2004; Nam *et al.*, 2012). The configurable articulatory synthesizer CASY (Iskarous *et al.*, 2003) is an embedded-model articulator that uses joint parameters and includes values from Mermelstein's original (1973) model-generated vocal tract articulation. More recently, *task-dynamic*-based models (Satzman & Munhall, 1989) have also utilized kinematic data from ariticulatory studies and registered estimated dynamic values for a model articulator. For instance, Alexander *et al.* (2017) applied the results of VC$_{vd}$V sequences from a real-time magnetic resonance imaging (rtMRI) experiment to a model articulator, and reconstructed a speaker-specific vocal tract. Regarding the bilabial stop, lip aperture is a tract variable, and the upper lip, lower lip, and jaw were likewise parameters of the model articulator in Alexander *et al.*'s (2017) study.

1.2. Vertical jaw movement

The jaw is composed of the *maxilla*, or upper jaw, and the *mandible*, or lower jaw. The *mandible* can move up and down when chewing and speaking. Vertical movement of the mandible is possible, as the *condyle* on the top of the *ramus* is connected to the temporal bone of the skull: the round-edged *condyle* and *mandibular fossa* of the temporal bone come together at the *temporomandibular* joint (see Gick *et al.* (2013:146) for a detailed review of the jaw).

In articulatory phonology, jaw height functionally serves to form varying constriction largely between consonants and vowels (Saltzman & Munhall, 1986; Browman & Goldstein, 1990). *Mandible* height (henceforth, *jaw* height) is further maneuvered in a complex manner. This differs for different vowels from Southern British English and Egyptian Arabic in an X-ray motion film study (Wood, 1979). German also showed that jaw height decreased in the order /u/ > /ʊ/ (Ladefoged & Maddieson (1996) after Bolla & Valaczkai (1986)). Jaw height in American English as produced by five speakers gradually decreased in the order /i/ > /ɪ/ > /ɛ/ > /æ/ for front vowels and in the order /u/ > /ʊ/ > /ɑ/ for back vowels (Ladefoged, 2001). Individual differences are observed for Gaelic vowels in terms of tongue height for two front vowels: one speaker gradually decreased tongue height in the order /ɪ/ > /e/, while the other speaker reversed the order to /e/ > /ɪ/, while jaw height remained consistent for both speakers in the order /ɪ/ > /e/ (Goldstein, USC class website).

For consonants, jaw height was affected by several factors such as voicing, speech style, place of articulation, and manner of articulation, but results varied across studies. Keating *et al.* 's (1994) cross-linguistic study examined jaw height of consonants in three homorganic VC'V contexts (/i/, /e/, /a/) using a movetrack magnetometer system. Jaw height gradually decreased but it was roughly divided into /s/, /t/, /d/, /r/, /f/ > /l/, /n/, /b/, /k/, /h/ for English and /s/, /t/, /d/, /f/, /n/, /r/ > /b/, /k/, /l/, /h/ for Swedish: overall, coronal obstruents (e.g., /s/, /t/, /d/) were consistently higher across the board. Examining voicing contrast (e.g., voiced vs. voiceless) and manner of articulation (e.g., stop vs. fricative vs. lateral), Mooshammer *et al.* (2007) conducted an electromagnetic midsagittal articulography experiment with five German speakers. In loud speech, some speakers demonstrated lower jaw height for coronal sonorants (/n/, /l/) compared to obstruents (/t/, /d/, /s/, /ʃ/): a coronal nasal (/n/) exhibited lower jaw height for four speakers out of five, and a lateral (/l/) for two speakers (loud > comfortable). Likewise, lower jaw position was observed for coronal nasal /n/ in comparison with coronal stops with varying laryngeal contrast (/t/, /tʰ/, /t*/) in a homorganic low-vowel context (/a/-to-/a/) for Korean (Son *et al.*, 2011). However, different speech rate effects were not empirically attested in Korean: lateral /l/ in homorganic intervocalic position (/...ala.../ and /...ili.../), from which flap /ɾ/ derives, exhibits similar jaw height in different speech rates (fast = comfortable) (Son, 2015a, 2015b).

1.3. Bilabial movement

Attention has been drawn to research on kinematic movement of articulators in various languages. Using various methodologies following the path of pellets' location at given points (e.g., the upper lip, lower lip, tongue tip, tongue body, etc.), bilabial constriction has been a matter of interest partly due to its unique physiological characteristics. That is, the upper lip moves downwards in coordination with the elevation of the lower lip paired with jaw raising, on the one hand, while the passive receding movement of the upper lip results from lower lip elevation (Gick *et al.*, 2013).

There are some previous studies which have provided kinematic movement data of the upper lip and the lower lip individually. Löfqvist's (1996) simultaneous two-dimensional magnetometer system (Perkell *et al.*, 1992) and air pressure study examined intervocalic bilabial stops in English (/apV/, /abV/) (/i/, /a/, /u/ in V) embedded in a carrier phrase (e.g., 'say ___ again'). The results from the production of three speakers of American English and

one Swedish speaker indicated that the lips move more after target attainment because the lip tissues are being compressed during acoustic silence. In Löfqvist (1993), similar results were also obtained due to the compression of the lips in a production study on lips movements, tongue body movements, and laryngeal movements, while simultaneously collecting articulatory, air pressure, and transillumination data from two subjects. In particular, an American-English speaker and a Japanese speaker showed the same pattern, indicating lip compression. As pointed out in Gick et al. (2013), target values for constricting articulators with a tight seal are negative (e.g., *overshoot*) so that speakers do not have to administer *fine control* to make constriction. In Son (2018), pellet locations for the upper and lower lips were traced in a two-dimensional magnetometer system (Perkell et al., 1992). Examining seven speakers of Seoul Korean producing the intervocalic voiceless stop /p/ in /a/-to-/a/ sequences within short natural sentences, she found that the upper lip moved further downwards as compensation for reduced lower lip raising movements with an intervening across-word boundary, compared to word internally.

In the investigation of bilabial stops, kinematic aspects of lip aperture have been systematically examined in terms of different syllabic position (e.g., onset vs. coda), prosodic contexts (e.g., pitch-accented vs. unaccented), and assimilating contexts (e.g., /t#k/ vs. /k#t/). For American English, Browman & Goldstein's (1995) microbeam study with one Californian male speaker showed that bilabial voiceless stop in the coda was more spatially reduced than the onset in terms of lip aperture of a target word *pop*, which was consistent across different prosodic contexts (e.g., post-pitch-accented position ('MY ***pop*** huddles); pitch-accented position (my ***'POP*** huddles); pre-pitch-accented position (my ***pop*** 'HUDDLES)). Using a two-dimensional magnetometer system (Perkell et al., 1992), Son (2008) examined lip aperture of /p/ in assimilating contexts (/ap(#)ka/) with five speakers of Seoul Korean, varying in speech rate and morphosyntactic conditions. She observed partial spatio-temporal reduction of lip aperture with a phrasal boundary in one speaker. In terms of different speech rates, spatiotemporal reduction was more frequent in fast rate than comfortable rate. Meanwhile, Son et al.'s (2007) kinematic study with three speakers of Seoul Korean showed gestural reduction of lip aperture in assimilating contexts, observing categorical reduction of the target /p/ if it ever occurred. In their study, categorical reduction was manifested in the within-word condition accompanied by more frequent occurrences in fast rate than comfortable rate.

1.4. Research questions

In this paper, we examine articulation of the jaw and the lips. In particular, we describe vertical jaw maxima and lip aperture (LA) minima. We also spell out relative timing relations between the jaw and the lips (upper and lower lips) in terms of their positional extremes.

Firstly, we aim to examine whether, and if so how, jaw height in the bilabial stop /p/ varies with either a linguistic factor, a morphosyntactic boundary (across-word vs. within-word), a paralinguistic factor, speech rate (comfortable vs. fast), or both. In articulatory phonology, the jaw has been assumed to serve a bifunctional purpose, namely consonantal constriction (more elevated) in contrast to vocalic constriction (more open) (Browman & Goldstein, 1990; Satzman & Munhall, 1986). Previous literature has shown that jaw height moves upwards, varying with place of

articulation, manner of articulation, or speech style (Keating et al., 1994; Mooshammer et al., 2007; Son, 2015a, 2015b; Son et al., 2011). Although results vary across studies, relatively higher jaw position was observed for coronal obstruents in a consistent way across studies (e.g., /t/, /d/, /s/, /ʃ/) while relatively lower jaw position was observed for non-coronal obstruents (e.g., /b/, /k/, /h/) (Keating et al., 1994). The coronal nasal /n/ consistently exhibited lower jaw position in loud speech. To quote Mooshammer et al. (2007:172), "... the lower jaw positions in loud speech during the nasal can be attributed to an accommodation of the jaw to the lower jaw positions in loud speech of the surrounding vowels." In other words, a coronal nasal /n/ is most likely to be influenced by surrounding vocalic articulation, being more sensitive to intergestural coarticulation. Coronal nasal /n/ also manifested lower jaw position in Seoul Korean, compared to its aspirated /tʰ/, fortis /t*/, and lenis /t/ counterparts (/n/</tʰ/; /n/≤/t*/=/t/ in Son et al., 2011), but different speech rates did not perturb jaw movements during the production of an intervocalic flap /ɾ/ derived from lateral approximant /l/ (Son, 2018a, 2018b). Since there have not been any studies which have rigorously explored whether a single segment systematically demonstrates different jaw height in terms of a linguistic (across-word vs. within-word) and/or paralinguistic (fast vs. comfortable) factor, we focus, in this paper, on the intervocalic bilabial stop /p/ in Seoul Korean. In this way, we will try to suggest finely tuning dynamic parameter values for jaw angle if it varies with a linguistic factor, a paralinguistic factor, or both.

Secondly, we examine lip aperture minima in terms of different word boundaries and speech rates. In terms of constriction of the lips in intervocalic stop consonants (/V$_1$pV$_2$/, /V$_1$bV$_2$/), Löfqvist (1996) found that the upper lip (UL) began receding upwards after reaching positional minimum values as it gave way up to the point in time when the lower lip (LL) raised its maximal point. This is attributed to compressing of the lips: as a result of *overshoot*; speakers do not have to make a constriction with [-continuant] with *fine control* (Gick et al., 2013). Notice that upper lip (UL) lowering occurred to compensate for the spatial reduction of the lower lip (LL) in the across-word boundary condition (Son, 2018). However, it was not obvious, from the perspective of coordinative lip constriction, whether the articulatory compensation occurred to the extent that it obliterated different word boundary effects (across-word = within-word) or simply prevented excessive lenition of /p/ to preserve word boundary effects intact (across-word < within-word). In this study, in an effort to determine different word boundary effects, we revisit intervocalic lip movement in terms of lip aperture minima (see also Alexander et al. (2017), Browman & Goldstein (1986, 1988, 1990, 1995), Kochetov et al. (2007), Ladefoged & Maddieson (1996), Löfqvist (1996), Löfqvist & Gracco (1997), Maddieson (2005), Smith (1992), and Son (2008) for various experimental methodologies collecting kinematic data of bilabial constriction). In this way, we aim to provide a comprehensive analysis of the intervocalic bilabial stop to resolve the two alternative interpretations as we suggested above.

Lastly, a further objective is to learn whether, and if so how, relative timing lags between the lips (upper and lower) and the jaw differ in terms of vertical positional extremes. Under the hypothesis of articulatory phonology, variability in casual speech including phonological processes and alternations is attributed to gestural overlap (Browman & Goldstein, 1990, 1991, 1992). Intergestural

timing has been relatively well studied cross-linguistically since it has served to provide empirical evidence for the gestural overlap-based hypothesis and to estimate dynamic parameter values for vocal tract constriction variables. In particular, consonantal clusters have been fairly well examined in this regard (/pt/, /tk/, /kt/, /kp/, /tjm/, /djb/ in Russian (Kochetov & Goldstein, 2005); /bg/, /pʰtʰ/, /dg/, /gb/, /tʰb/, /gd/ in Georgian (Chitoran et al., 2002); /t#k/ in British English (Nolan, 1992); /pk/, /p#k/ in Seoul Korean (Son et al., 2007); /d#k/, /g#k/, /d#h/ in English and /t#k/, /k#k/, /t#h/ in German (Kühnert & Hoole, 2004); /ks/, /kt/, /pt/ in Seoul Korean (Son, 2013); inter alia). In addition, single segments have been examined as in a research focus like intergestural timing of active vocal tract variables to account for different syllable positional effects (e.g., *leap* [onset] vs. *peel* [coda] in Browman & Goldstein (1995)). In particular, horizontal tongue body retraction and vertical tongue tip movement were simultaneously coordinated for the onset but sequentially for the coda (see also *more* [onset] vs. *seem* [coda] with respect to the lips and the velic opening gestures in Krakow (1989)). In this paper, our analytical focus narrows down on interarticulator relative timing between the jaw and the lips in terms of different phrase boundaries (across-word vs. within-word) and different speech rates (comfortable vs. fast) in Seoul Korean. In this way, we aim to improve our understanding of coordinative temporal movement of the three participating articulators involved in bilabial constriction (jaw angle (JA), upper lip height (UH), and lower lip height (LH)).

2. Method

2.1. Participants, data collection, and measurement

We revisited kinematic data used in Son (2018), which was acquired by using the two-dimensional point-tracking system, electromagnetic midsagittal articulometer (EMMA in Perkell et al., 1992). It was originally collected from seven (four female and three male) native speakers of Seoul Korean (Seoul or Gyeonggi province in South Korea) who were in their mid-twenties and early thirties. They resided in Connecticut, U.S.A. when they participated in the production experiment and received a financial remuneration for their participation[1].

Kinematic data was mathematically expressed as a vector on an ordinate plane. At the time of kinematic data collection, acoustic data was also obtained simultaneously. We used the positional values of electric transducers (i.e., pellets) attached to three articulators: the upper lip, the lower lip, and the lower incisor (as an index of jaw movement) for further analysis. In particular, jaw maxima was measured within the time span of activation duration of lip aperture constriction. We reproduced the stimuli list in (1), borrowed from Son (2018:26) (see Son (2018) for a detailed description of the elicitation methodology used). A total of 223 tokens from seven speakers were analyzed (7 speakers × 2 boundaries × 2 speech rates × 8 repetitions), with one one token being omitted due to stuttering.

(1) Stimuli (reproduced from Son (2018:26))
a. Target sequence /pa/
i. Within-word boundary condition
/apai/ 'father' (North Korean dialect)

ii. Across-word boundary condition
/pakatʃi/ '(a) gourd dipper'

b. Natural short sentence including the target sequence, its syntactic structure, and a symbol for a word boundary (#)

i. Within-word boundary condition
/apai # toŋmunɨn # pukhanmalija/
[IP[NP apai toŋmunɨn] [VP[NP pukhanmal] [V ija]]]
'Father comrade is North Korean vocabulary.'

ii. Across-word boundary condition
/tʃəna # pakatʃilɨl # pʰala/
[IP[N P tʃəna] [VP[NP pakatʃilɨl] [V pʰala]]]
'Jeona sells gourd dippers.'

MVIEW (Tiede, 2005) is software for analyzing kinematic data of articulation relevant to human speech. We used the function of *lp_Snapex* in MVIEW to determine maximum vertical jaw position, maximum vertical lower lip position, and minimum vertical upper lip position of the bilabial stop /p/ (see Son (2018) for details of gestural demarcation). Figure 1 illustrates the specifics of positional extremes of the three articulators, duplicated and captured from the temporal display in MVIEW. Precise time points are superimposed on four identical real-time movement trajectories in Figures 3.a.i, 3.b.ii, 3.c.iii, & 3.d.iv.

2.2. Statistical analysis

We converted raw data to z-scores before our statistical analysis was conducted. Linear mixed-effects models in R (R Development Core Team, 2014) were used for data analysis. The results of articulatory analysis in z-scores were fitted with the lemr function from the lem4 packages (Bates et al., 2011). In particular, we fitted a linear regression model on maximum vertical jaw position, lip aperture, minimum vertical upper lip position, and maximum vertical lower lip position as we looked into Speech rate (fast vs. comfortable) and Boundary (across-word vs. within-word), with Subject as the random intercept[2]. We conducted likelihood ratio tests using ANOVA (analysis of variance) in order to evaluate interactions (full model and interaction model) and main effects (null model and full model).

---

2 Twenty-seven tokens were excluded from statistical analysis since the highest position of the jaw (during /p/) were not found using *lp-Snapex*.
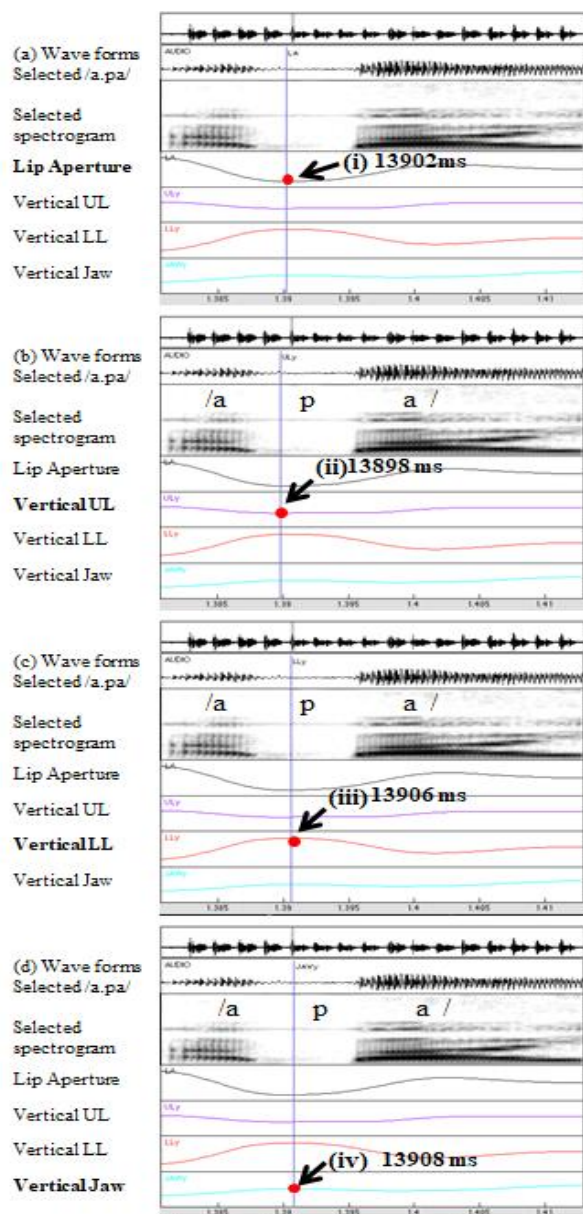
Figure 3. (a) describes lip aperture minima in (i). (b) describes the minimum vertical upper lip in (ii). (c) describes the maximum vertical lower lip in (iii). (d) describes the maximum vertical jaw in (iv). Greater values signify higher position for the articulatory movement of the upper lip, the lower lip, and the jaw. Also shown in Figures a.i, bii, c.iii, d.iv are time points where the minimum/maximum value of each kinematic trajectory is specified. The captured window depicts the bilabial stop /p/ in the context /a/-C-/a/, which was duplicated from the first token of the within-word boundary condition at the comfortable rate as spoken by a female speaker (SF1).

3. Results

3.1. Vertical jaw position

There was no interaction between Speech rate and Boundary (χ^2(1)=3.04, p>0.05). The results indicated that maximum vertical jaw position varied with Boundary (χ^2(1)=15.76, p<0.0001), but not with Speech rate (χ^2(1)= 3.11, p>0.05).

In Table 2, the results of linear mixed-effects models are shown in terms of vertical jaw maxima. Vertical jaw position exhibited

similar values in terms of Speech rate (t(189.07)=−1.77, p>0.05) (comfortable = fast) as shown in Figure 4.a. In terms of different boundary types (across-word vs. within-word), the within-word condition showed greater jaw height, raised by 1.06 mm (SE ±0.26) (t(189.3)=4.06, p<0.0001) (across-word < within-word), as shown in Figure 4.b.

Table 2. The results of linear mixed effects models on vertical jaw maxima

	Estimate	SE	df	t value	Pr(>|t|)
(intercept)	-2.41	0.77	7.79	-3.14	p<0.05
Speech rate [fast]	-0.46	0.26	189.07	-1.77	p>0.05
Boundary[within-word]	1.06	0.26	189.30	4.06	p<0.0001

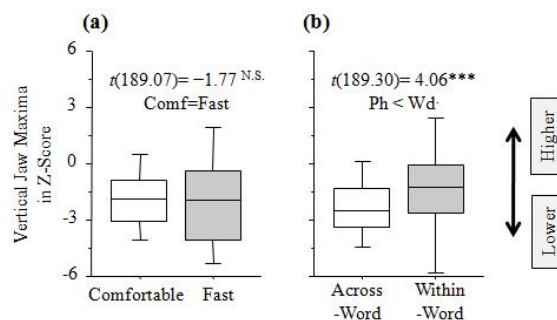*Note*: Number of observations: 196. Groups: subject, 7.

Figure 4. Vertical jaw maxima measured with (a) Speech rate and (b) Boundary. Greater values represent higher jaw position. (The symbol '***' is for p<0.0001.)

3.2. Lip aperture

The results of linear mixed-effects models showed that lip aperture (LA) varied with Speech rate (χ^2(1)=5.34, p<0.05), but not with Boundary (χ^2(1)=0.03, p>0.05). There was significant interaction between Speech rate and Boundary (χ^2(1)=4.11, p<0.05). However, the results of t-tests did not render statistical significance in terms of different word boundaries (t(84.70)=−1.61, p>0.05 (across-word = within-word) for fast rate; t(104.87)=1.22, p>0.05 (across-word = within-word) for comfortable rate).

In Table 3, lip aperture exhibited more reduction in fast speech rate, with a reduction to 0.32 mm (SE ±0.14) (t(196)=2.33, p<0.05) (comfortable < fast) as shown in Figure 3.a. In terms of different boundary types, lip aperture did not change (t(196)=−0.18, p>0.05) (across-word = within-word), as shown in Figure 3.b

Table 3. The results of linear mixed effects models on lip aperture

	Estimate	SE	df	t value	Pr(>|t|)
(intercept)	-0.13	0.12	196	-1.13	p>0.05
Speech rate [fast]	0.32	0.14	196	2.33	p<0.05
Boundary[within-word]	-0.03	0.14	196	-0.18	p>0.05

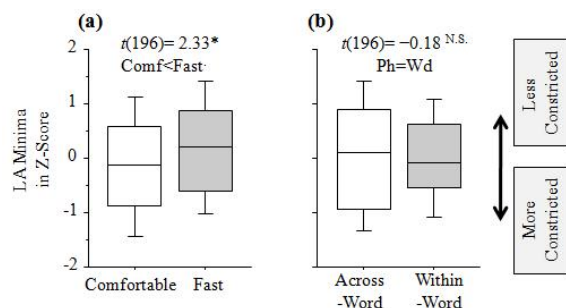*Note*: Number of observations: 196. Groups: subject, 7.

Figure 5. Lip aperture (LA) minima measured with (a) Speech rate and (b) Boundary. Greater values represent less constriction (i.e., more reduction). (The symbol '*' is for $p<0.05$.)
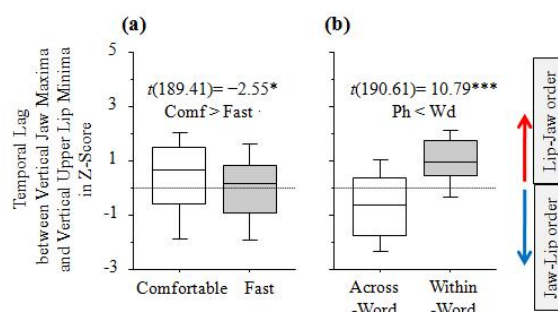
3.3. Temporal lag between vertical jaw maxima and vertical lip extremes

3.3.1. Temporal lag between vertical jaw maxima and vertical upper lip minima

There was no interaction between Speech rate and Boundary ($\chi^2(1)=2.08$, $p>0.05$). The results indicated that the temporal lag between the time point of the maximum vertical jaw position and that of the minimum vertical upper lip (UL) position varied with Speech rate ($\chi^2(1)=6.37$, $p<0.05$) as well as Boundary ($\chi^2(1)=89.15$, $p<0.0001$).

In Table 4, the results of linear mixed-effects models are shown in terms of the temporal lag between the time point of vertical jaw maxima and that of vertical upper lip (UL) minima. For different speech rates, we observed that the time point of maximum vertical jaw position preceded that of minimum vertical upper lip (UL) position, advancing the time point of minimum vertical upper lip (UL) position by −0.40 ms in fast rate, as compared to comfortable rate ($SE \pm 0.16$) ($t(189.41)=−2.55$, $p<0.05$) (comfortable > fast). As a result, the temporal lag between the time point of vertical jaw maxima and that of vertical upper lip (UL) minima in fast rate is characterized by synchrony, being more approximated to zero as shown in Figure 6.a. In terms of Boundary effects, the time point of the maximum vertical jaw position preceded that of the minimum vertical upper lip (UL) position in the across-word condition (e.g., *negative values* in this case), while the reverse order was observed in the within-word condition (e.g., *positive values* in this case) as shown in Figure 6.b. In particular, the time point of vertical jaw maxima occurred later than that of upper lip (UL) minima in the within-word condition, with a lag of 1.69 ms ($SE \pm 0.16$) ($t(190.61)=10.79$, $p<0.0001$), as compared to the across-word condition (across-word < within-word).

Table 4. The results of linear mixed-effects models on temporal lag between vertical jaw maxima and vertical upper lip minima

	Estimate	SE	df	t value	Pr(>|t|)
(intercept)	-0.47	0.22	11.92	-2.17	*p*>0.05
Speech rate [fast]	-0.40	0.16	189.41	-2.55	*p*<0.05
Boundary[within-word]	1.69	0.16	190.61	10.79	0<0001

Note: Number of observations: 196. Groups: subject, 7.

Figure 6. Temporal lag between the time point of vertical jaw maxima and that of vertical upper lip (UL) minima with (a) Speech rate and (b) Boundary.
Negative values indicate that the time point of vertical jaw maxima precedes that of vertical upper lip (UL) minima. Positive values indicate that the time point of vertical jaw maxima follows that of vertical upper lip (UL) minima. Zero values indicate that the time point of vertical jaw maxima occurs in synchrony with that of vertical upper lip (UL) minima. (The symbols '*', '***' are for $p<0.05$, $p<0.0001$, respectively.)

3.3.2. Temporal lag between vertical jaw maxima and vertical lower lip maxima

There was no interaction between Speech rate and Boundary ($\chi^2(1)=1.06$, $p>0.05$). The results indicated that the temporal lag between the time point of the maximum vertical jaw position and that of the maximum vertical lower lip (LL) position varied with Speech rate ($\chi^2(1)=6.56$, $p<0.05$) and Boundary ($\chi^2(1)=86.50$, $p<0.0001$).

In Table 5, the results of the linear mixed-effects models are shown. In terms of Speech rate, the time point of vertical jaw maxima followed that of vertical lower lip (LL) maxima, reducing temporal lag by −0.34 ms ($SE \pm 0.13$) ($t(189.32)=−2.58$, $p<0.05$) (comfortable > fast) as shown in Figure 7.a. The median value of the fast speech rate was also more approximated to zero, indicating a synchronous coordination between two articulatory events in terms of positional maxima. In terms of Boundary effects, the time point of the maximum vertical jaw position followed that of the maximum vertical lower lip (LL) position in the within-word condition (e.g., *positive values* in this case), while the reverse order was true in the across-word condition, as shown in Figure 7.b. The time point of vertical jaw maxima occurred later on that of lower lip (LL) maxima in the within-word condition, with a lag of 1.40 ms ($SE \pm 0.13$) ($t(190.28)=10.62$, $p<0.0001$), as compared to the across-word condition (across-word < within-word). To conclude, temporal lags between vertical jaw maxima and lip constriction extremes (UL and LL, individually) exhibited similar results.

Table 5. The results of linear mixed-effects models on temporal lag between vertical jaw maxima and vertical lower lip maxima

	Estimate	SE	df	t value	Pr(>|t|)
(intercept)	-0.37	0.20	10.75	-1.83	*p*>0.05
Speech rate [fast]	-0.34	0.13	189.32	-2.58	*p*<0.05
Boundary[within-word]	1.40	0.13	190.28	10.62	0<0.0001

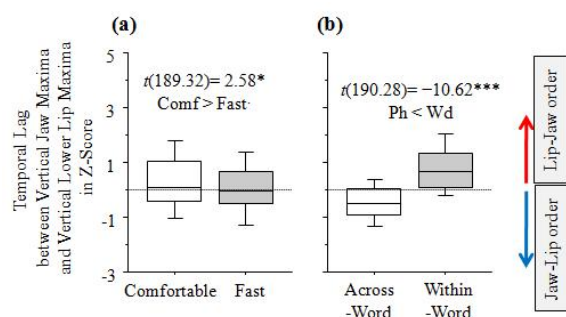*Note*: Number of observations: 196. Groups: subject, 7.

Figure 7. Temporal lag between the time point of vertical jaw maxima and that of vertical lower lip (LL) maxima with (a) Speech rate and (b) Boundary.
Negative values indicate that the time point of vertical jaw maxima precedes that of vertical lower lip (LL) maxima. Positive values indicate that the time point of vertical jaw maxima follows that of vertical lower lip (LL) maxima. Zero values indicate that the time point of vertical jaw maxima occurs in synchrony with that of vertical lower lip (LL) maxima. (The symbols '*', '***' are for $p<0.05$, $p<0.0001$, respectively.)

4. Summary and Discussion

We addressed how relevant articulators are coordinated in terms of spatiotemporal aspects of the bilabial stop /p/ in the intervocalic context /a/-to-/a/ in Seoul Korean. Firstly, the effects of linguistic (different word boundary) and paralinguistic (different speech rate) factors on vertical jaw maxima in the bilabial stop /p/ were examined, and we found out that this only varied with different word boundary conditions, not with different speech rates. In particular, vertical jaw maxima demonstrated lower jaw position in the across-word boundary condition, indicating that the jaw contributes less to bilabial constriction in this context. With different jaw position dependent on morphosyntactic boundaries, lip aperture was similar between two different word boundary conditions. In contrast, rate-dependent variation exhibited the opposite pattern in terms of vertical jaw position and lip aperture — lip aperture varied with different speech rates (e.g., less constriction in fast rate) while vertical jaw position did not. Secondly, relative timing relations between vertical jaw maxima and vertical upper lip (UL) minima varied with different word boundaries as well as different speech rates. With regard to different speech rates, a more synchronous relation was observed in fast rate. In terms of different word boundaries, vertical jaw maxima preceded vertical upper lip (UL) minima in the across-word boundary condition, but occurred after vertical upper lip (UL) minima word-internally. Likewise, a similar pattern was observed with vertical lower lip (LL) maxima. Combining the vertical jaw maxima values with the relative timing relations between vertical jaw maxima and lips constriction extremes, the jaw stopped moving upwards before bilabial constriction reached its maximal position in /...a#pa.../ sequences. This could have led to lower jaw position in the across-word boundary condition.

4.1. Lower jaw position in the across-word boundary condition and the phrase boundary condition

Previous literature on varying jaw position has concentrated on segmental contrasts in terms of vocalic articulation, place of articulation, and manner of articulation (Browman & Goldstein, 1990; Keating *et al.*, 1994; Ladefoged, 2001; Ladefoged & Maddieson,

1996 after Bolla & Valaczkai, 1986; Mooshammer *et al.*, 2007; Son *et al.*, 2011; Wood, 1979; *inter alia*). Although results vary across studies, the findings relating to jaw height can be generalized such that i) consonants are associated with higher jaw position in comparison with vowels (Browman & Goldstein, 1990; Wood, 1979) and ii) consonants are classified in the order coronal>labial= dorsal. Among coronal consonants, lower jaw position was manifested in loud speech for nasal coronal /n/ or lateral approximant /l/, as compared to coronal obstruents (/s/, /t/, /d/, /f/) and coronal approximant (/r/) for English and Swedish (Mooshammer *et al.*, 2007). Meanwhile, jaw height differed among coronal consonants as a function of manner of articulation in Korean, demonstrating lower jaw position in coronal nasal /n/ as compared to coronal stop with ternary laryngeal contrasts (/n/</tʰ/; /n/≤/t*/=/t/ in Son *et al.* (2011)). The current study further showed that jaw height even varied within a single segment, /p/, as a function of different word boundaries, with reduction in the across-word boundary condition.

Vertical jaw movement is relatively free to maneuver on a continuum without segmental confusion incurred, compared to primary constrictors such as the tongue body and the tongue tip (Nam, ms.). According to Nam's explanation of coordinative movements for speech actions, articulator position at a given time can be described using a vector represented by x and y on an ordinate system. For the position of the tongue body at a given time, three factors are employed in principle - the jaw angle (JA), the tongue body center angle (CA), and the distance from the condyle and the center of the body ($\overline{FC}$) (see Mermelstein's model articulator in Figure 1). He pointed out that a problem arises due to the excessive number of factors (e.g., JA, CA, $\overline{FC}$) for an observed value with two degrees of freedom, x and y. This, in turn, enables speakers to be considerably free with using the jaw for tongue body movements, which generates interspeaker or intraspeaker variability in terms of vertical maneuvering of the jaw articulator. In applying this mechanism to the findings from the current study, a time-dependent vector (e.g., lip aperture (LA) as a tract variable) involves three degrees of freedom (i.e., the upper lip height (UH), the lower lip height (LH), and the jaw angle (JA)) for making a lip aperture gesture in a computational model. Under this mechanism, bilabial constriction can be completed without the assistance of jaw raising movement since it is considered more than necessary. As a possibility, the results of the current study on morphosyntactic-boundary-dependent jaw movement can be reflected in the level of model articulator in a computational model (i.e., jaw angle) by specifying dynamic parameters accordingly (i.e., greater weight), such that it can suppress vertical jaw movement in the across-word boundary condition.

A reviewer pointed out that the results of vertical jaw movement might generally benefit by considering the prosodic structure of Seoul Korean, expressing the concern that morphosyntactic boundary conditions (across-word vs. within-word) can occur at the prosodic level and the prosodic structure (a linguistic factor) of an utterance can be influenced by speech rate (a paralinguistic factor). In particular, it was pointed out that labial in the across-word boundary condition (/tʃəna # **p**akatʃi/) could be produced at the edge of a phrase (IP-initial or AP-initial) at comfortable rate as opposed to phrase-internally (IP-internal or AP-internal) at fast rate. In response to this concern, we examined pitch contour during the production of the sequence /tʃəna$_{V1}$ # pa$_{V2}$ka$_{V3}$tʃiłɨl # pala/ in across-word boundary condition. f0 measurements were extracted at

three time points, i.e., i) at the end point of V1 (/a/), ii) during V2 (/pa/), taking the average f0, and iii) during V3 (/tʃi/), taking the average f0. Among several possible prosodic readings for the first two words /tʃəna$_{V1}$ # pa$_{V2}$ka$_{V3}$tʃilil/, we considered two different prosodic phrasings following Jun (1993). One was {$_{AP}$ tʃəna} {$_{AP}$ pakatʃilil} and the other {$_{AP}$ tʃəna pakatʃilil} (note that we did not take into account how the sequence /pala/ is prosodically grouped for the sake of simplicity of analysis and the symbol '{ }' represents an accentual phrase demarcation). The tonal pattern of an accentual phrase is T(HL)H where a H(igh) tone is assigned to T if a phrase-initial phoneme is [+stiff vocal folds] (Halle & Stevens, 1971; Jun, 1993); otherwise, a L(ow) tone is assigned. We would expect a rising contour along /pa/ in V2 and /ka/ in V3 if an accentual phrase boundary coincides with a word boundary, but a falling contour if V2 and V3 belong to the preceding word /tʃəna/ comprising one accentual phrase. Subtracting the f0 value of V2 from that of V3, a positive value denotes a rising contour, which in turn brings about two accentual phrases during production (e.g., {$_{AP}$ tʃəna} {$_{AP}$ pakatʃilil}); otherwise, a falling contour emerges to indicate one accentual phrase (e.g., {$_{AP}$ tʃəna pakatʃilil}) with a falling contour over V2V3 which is a byproduct of an interpolation between H in /na/ and L in /tʃi/. 64% of tokens were produced with two accentual phrases in the across-word boundary condition (71 out of 111 tokens) and 34 % of tokens within one accentual phrase (38 out of 111 tokens). Two tokens were excluded from further analysis since f0 change had not been detected in them. Perceptual judgments have not considered in analysis.
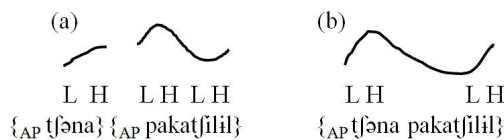

Figure 8. Schematic representation of tonal contours (a) {$_{AP}$ tʃəna} {$_{AP}$ pakatʃilil} in two accentual phrases and (b) {$_{AP}$ tʃəna pakatʃilil} in one accentual phrase. Allophonic variation involving voicing and flapping is not transcribed.

Examining a subset of data where vertical jaw maxima measurements were available (i.e., 95 tokens in the across-word boundary condition), we observed that 24 tokens produced at fast rate were yielded AP-internally and 20 tokens AP-initially. At comfortable rate, 7 tokens were produced AP-internally while 44 tokens AP-initially; more tokens at comfortable rate were produced with AP-initial position, that is, a phrasal boundary location. We fitted a linear regression model on maximum vertical jaw position as we looked into Speech rate (fast vs. comfortable) and Prosody (AP-initial vs. AP-internal), with Subject as the random intercept. The results of maximum vertical jaw position across word boundaries showed that there was no interaction (χ^2(1)=1.03, p>0.05). The dependent variable varied with Speech rate, but not with Prosody. The results of t-tests rendered statistical significance in terms of Speech rate, lowered by −0.78 mm in fast rate (*SE* ±0.26) (t(89.88)=−3.05, p<0.01) (comfortable > fast), but not with Prosody (t(91.88)=1.74, p>0.05 (AP-boundary = AP-internal)). We further examined each speech rate condition separately, fitting a linear regression model on maximum vertical jaw position as we looked into Prosody (AP-initial vs. AP-internal). For fast speech rate, the results of t-tests rendered statistical significance in terms of

Prosody, raised by 0.69 mm AP-internally (*SE* ±0.22) (t(37.91)= 3.14, p<0.001) (AP-initial < AP-internal), while no statistical significance was observed for comfortable speech rates in terms of Prosody (t(45.62)=−1.31, p>0.05) (AP-initial = AP-internal). None of the results indicated that AP-initial position exhibited higher jaw position.

Previous literature of prosodically driven articulation or co-articulation has shown that higher prosodic domains are generally associated with articulatory strengthening and less coarticulation (Cho, 2004; Cho & Keating, 2001, 2009; Cho *et al*., 2016; Keating *et al*., 2003; *inter alia*). Linguopalatal contact was greater in higher prosodic domains (e.g., Utterance-initial (Ui) and Intonational phrase-initial (IPi)) than in lower prosodic domains (e.g., Accentual phrase-initial (APi) and Word-initial (Wi)), (Ui, IPi > APi, Wi), although there was interspeaker variability. In some measurements (e.g., linguopalatal contact and seal duration), APi, with interspeaker variability in terms of linguopalatal contact, was not a prosodic condition for domain-initial strengthening, as compared to Wi (Keating *et al*., 2003). Domain-initial strengthening (Ui, IPi vs. U-internal, IP-internal) is not locally restricted to boundary-initial consonants (/n/, /t/), but is also globally attested with the vowel (/ɛ/) in the examination of CVs, where the target syllable appeared in a trisyllabic word (/nɛbəbɛn/, /tɛbəbɛn/) embedded in carrier phrases ('___ fed them.' and 'one deaf ___') (Cho & Keating, 2009). Pondering over the idiosyncratic results observed with our vertical jaw maxima data, we turn to Öhman's (1966) finding in which a consonant is superimposed onto the consecutive vocalic lingual movement from V1 to V2. In particular, vocalic lingual articulation is physiologically unconstrained by an intervening labial consonant (e.g., /p/), compared to a lingual consonant (e.g., /t/ or /k/) (Kühnert, 2006; Öhman, 1966; Recasens, 1984). We began with solving a conundrum involving vocalic articulation by giving more weight to articulatory strengthening of adjacent vowels in stronger prosodic locations, and now gear into the idiosyncratic behavior of bilabial stop /p/ in terms of jaw position (i.e., lower jaw position in the across-word boundary condition and in the AP-initial position).

Notice that the jaw articulator is repeatedly used in consonantal articulation as well as vocalic articulation (see Table 1). Most previous studies have examined primary articulation (e.g., the lips for labial, the tongue tip for coronal, etc.), while our data also showed reduction in participating articulators during the production of bilabial stop /p/ in terms of maximum vertical jaw position. As pointed out earlier (see section 1.1 for a detail), /p/ involves tract variables LA and LP/PRO while the jaw articulator (JA in this case) is one of the joint variables engaged in these tract variables. Given that the strengthening of vocalic articulation is manifested by a more open vocal tract with greater jaw lowering in low vowel /a/, this may have rendered more reduction in consonantal articulation (vertical jaw position in this case). Since our data include /p/ in the homorganic low vowel /a/-to-/a/ context, it is plausible that domain-initial adjacent low vowels ({$_{AP}$ tʃəna} {$_{AP}$ pakatʃilil}) may demonstrate lower jaw position, which may have in turn influenced the intervening consonant /p/ to the extent that it exhibits lower jaw position, being assimilated to lower jaw position in the /a/-to-/a/ context, in higher prosodic domains (e.g., AP-initial < AP-internal). From this perspective, reduction in jaw height in an intervening consonant in the /a/-to-/a/ context can possibly be understood as a byproduct of articulatory strengthening of vowels (i.e., lower jaw position) in prosodically stronger locations. In exploring the

coordinative movement of articulators engaged in the lip aperture gesture of intervocalic bilabial stop /p/, we have found empirical evidence that at a minimum, the jaw articulator should be carefully investigated along with a tract variable, LA, for a bilabial stop (e.g., /p/) in articulatory studies, so that we can enhance our understanding of participating articulators structured coordinatively at the segment level as well as any relation between jaw movement patterns and prosody.

To conclude this section, some caution should be taken since prosodic analysis is quite limited and incomplete in the current study and needs to be further analyzed to figure out exactly what occurred in speech articulation (e.g., with regard to the target consonant as well as adjacent vowels). In future study, we also need to address possibilities to explain how the prosodic structure (a linguistic factor) of an utterance can be influenced by speech rate (a paralinguistic factor) in a systematic way in terms of the tract variables and articulators involved.

4.2. Speech-rate effects applied to a computation model

4.2.1. Spatial reduction in fast speech rate

The current study showed that the lip aperture gesture was more reduced in fast rate, compared to comfortable rate. In the *task-dynamic* model of speech production (e.g., *task dynamics application* toolkit, TADA, in Nam *et al*. (2012)), speech rate-ependent gestural reduction is controlled at the gestural score level. Activation time values entered for an active tract variable are inherently smaller for fast rate, therefore target attainment cannot be completed simply due to shortness (or a lack) of time, and spatial reduction in lip aperture occurs as a consequence.

4.2.2. More synchronous temporal relation between the jaw and the upper and lower lips in fast speech rate

Vertical jaw position in the bilabial stop /p/ did not spatially vary with different speech rates, being compatible with the results of previous articulatory studies on derived intervocalic flaps (comfortable = fast) (e.g., [aɾa] in Son (2015a) and [iɾi] in Son (2015b)). In contrast, speech rate was a factor to differentiate relative temporal relations between the jaw and the lips, which were characterized by a synchronous coordination in fast rate. To quote Byrd (1996:139), "A variety of work has demonstrated that articulatory, prosodic, and extralinguistic factors all influence speech timing in a complex and interactive way." Speech timing has been investigated in terms of intergestural timing in a variety of articulatory studies (Browman & Goldstein, 2000; Nam, 2007; Nam *et al*., (in press); Saltzman *et al*., 2006). In particular, properties relating to the syllable structure (e.g., onset vs. coda) of an utterance are distinctly represented: onset is specified with an in-phase (0°) relation and coda an anti-phase (180°) relation in the coupling graph (Nam, 2007). Saltzman *et al*.'s (2006) theoretical basis was grounded in Haken, Kelso, & Bunz (1985) where human hand movements abruptly became more synchronous with increasing rate, from anti-phase (unstable mode) to in-phase (stable mode) relations. Note that a model articulator such as TADA specifies phase relations at the level of intergestural coordination (e.g., synchronous C-V in the onset vs. sequential V-C in the coda (Nam, 2007); synchronous tongue tip raising with respect to tongue body retraction in the onset vs. sequential relationship in the coda (Browman & Goldstein, 1995)). In line with this, the results of the current study

may suggest (or support) that a model articulator include a way to add phase relations to speech articulators involved in a segment observed in human speech (e.g., more stable mode of interarticulator locking in fast speech). Future studies on articulatory robotics should include this kind of issue, if possible.

4.3. Sptiaotemporal interarticulatory coordinative movement manifested in the upper lip, lower lip, and jaw

In the across-word boundary condition, the temporal point of maximum vertical jaw position preceded that of the upper lip. This can be understood as a premature termination of the assistance of the jaw in terms of assisting lower lip raising movement and lip tissue compression.

Son (2018) already showed more reduction of the lower lip movements of the intervocalic bilabial stop /p/ in the across-word boundary condition. Examining the same set of stimuli from that study in an effort to provide an understanding of the articulation of bilabial stop /p/, the current study also showed that vertical jaw position was lower in that particular context. Since the jaw and the lower lip are physiologically bound to one another (Gick *et al*., 2013), paired articulatory reduction may result in. Due to this physiological binding, there are two possible explanations for why reduction of vertical lower lip movement arises. One explanation is that spatial reduction in the vertical movement of the lower lip (LL) could have induced spatial reduction of jaw elevation. If we suppose that articulatory reduction of the lower lip (LL) could have induced spatial reduction of the jaw for /p/, this would suggest that spatial reduction of the lower lip (LL) elevation had anticipated spatial reduction of the jaw, and lip compression occurred subsequently without the assistance of the jaw. The other explanation is that spatial reduction of vertical jaw elevation could have caused reduction of raising movement of the lower lip. In line with this, we suppose that spatial reduction of the jaw could have induced that of the lower lip (LL) in the across-word boundary condition and the subsequent lip compression can be understood as the further independent raising movement of the lower lip as an effort to avoid lenition. Lower jaw position in the across-word boundary condition can be analogous to natural jaw yanking from vocalic gestures, which could have acted upon the concurrent lower lip reduction (see robotic jaw yanking by force with human subjects in Shiller *et al*. (2005)).

In Son (2018), spatial reduction of lower lip (LL) raising was resolved by compensatory upper lip (UL) lowering. Combining the results of the lower lip and the upper lip from the perspective of an articulatory task to complete a tight seal and release for labial stop /p/, we construe that articulatory compensation could have occurred in order to avoid lenition (e.g., intervocalic bilabial stop /p/ to labial approximant /w/). It is possible that articulatory compensation could have arisen to the extent of annihilating word-boundary effects so that a lip-closing gesture could have occurred in a coordinative articulatory effort in this particular context. By conducting an additional analysis of lip aperture (LA) minima in the current study, we were able to verify that labial constriction did not vary with different word boundaries, fitting linear mixed-effects models on lip aperture data relating to the number of observations in Son's (2018) study ($\chi^2(1)= 3.79$, $p \leq 0.05$ for interaction; $\chi^2(1)=7.98$, $p<0.01$ ($t(223)=2.85$, $p<0.01$) for speech rate (comfortable < fast); $\chi^2(1)=0.58$, $p>0.05$ ($t(223)=-0.76$, $p>0.05$) for boundary (across- word = within-word)). The result is

compatible with the current lip aperture data that was selectively analyzed to be in balance with available vertical jaw maxima (number of observations (196 tokens) as shown in Table 3), supporting compensatory lowering by the upper lip, possibly to avoid lenited [w]. One piece of supporting evidence can be found in a mechanical jaw perturbation experiment using a robotic jaw-yanking device in Shiller *et al.* (2005). In their study, speakers reacted voluntarily to arbitrary jaw perturbations such that they increased jaw stiffness in vowel production. In reaction to lowered jaw position, our subjects made use of another contributing articulator (the upper lip (UL) in this case) in a functional way such that articulatory compensation occurred by means of increasing upper lip (UL) lowering movement. We conclude that speakers of Seoul Korean have a holistic knowledge in producing speech from the perspective of *task*-based achievement by employing contributory vocal tract articulators in a functional way.

Acknowledgement

References

Alexander, R., Sorensen, T., Toutios, A., & Narayanan, S. (2017). VCV synthesis using task dynamics to animate a factor-based articulatory model. *Proceedings of Interspeech*. 244-248.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version, 1, 1-23. Retrieved from http://CRAN.R-project.org/package=lme4 on August 30, 2018.

Bolla, K., & Valaczkai, I. (1986). *Német Beszédhangok Atlasza (A Phonetic Conspectus of German)*. (Magyar Fonetikai Füzetek (Hungarian Papers in Phonetics) 16). Hungarian Academy of Sciences, Budapest.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45, 140-155.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.

Browman, C. P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299-320.

Browman, C. P., & Goldstein, L. (1991). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech* (pp. 341-376). Cambridge, U.K.: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Browman, C. P., & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-berti & L. Raphael (Eds.), *Producing speech: Contemporary issues*. For Catherine Safford Harris (pp. 19-33). AIP Press: New York.

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, 5, 25-34.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24, 209-244.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian, in C. Gussenhoven and Warner (Eds.), *Papers in Laboratory Phonology VII* (pp. 419-448). Berlin: Mouton de Gruyter.

Cho, T. (2004). Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32, 141-176.

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155-190.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37, 466-485.

Cho, T., Son, M., & Kim, S. (2016). Articulatory reflexes of the three-way contrast in labial stops and kinematic evidence for domain-initial strengthening in Korean. *Journal of the International Phonetic Association*, 46, 129-155.

Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory phonetics*. Oxford : Wiley-Blackwell.

Goldstein, L. USC class website. Retrieved from https://sail.usc.edu/~lgoldste/General_Phonetics/Slides/Vowels.pdf on October 3, 2018.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51, 347-356.

Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *Quarterly Progress Report* (Research Laboratory of Electronics, MIT) 101, 198-212.

Iskarous, K., Goldstein, L., Whalen, D., Tiede, M., & Rubin, P. (2003). CASY: The Haskins configurable articulatory synthesizer. *Proceedings of the 15th International Congress of Phonetic Sciences*, 185-188.

Jun, S-A. (1993). *The phonetics and phonology of Korean prosody*. Ph.D. Dissertation. Ohio State University.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In John Local, Richard Ogden, and Rosalind Temple (Eds.), *Papers in Laboratory Phonology VI* (pp. 145-163). Cambridge: Cambridge University Press.

Keating, P., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, 22, 407-422.

Kochetov, A., & Goldstein, L. (2005). Position and place effects in Russian word-initial and word-medial clusters, *Journal of the Acoustical Society of Merica*, 117, 2571.

Kochetov, A., Pouplier, M., & Son, M. (2007). Cross-language differences in overlap and assimilation patterns in Korean and Russian. *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1361-1364). Saarbrücken, Germany.

Krakow, R. (1989). *The articulatory organization of syllables: a kinematic analysis of labial and velar gestures*. Ph.D. Dissertation. Yale University.

Kühnert, B., & Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinical Linguistics and Phonetics*, 18, 559-575.

Kühnert, B., Hoole, P. & Mooshammer, C. (2006). Gestural overlap and c-center in selected French consonant clusters. *Proceedings of 7th International Seminar on Speech Production* (ISSP) (pp.

327-334), Brazil.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. Malden: Blackwell publishing Ltd.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Löfqvist, A. (1993). Electromagnetic transduction techniques in the study of speech motor control. *PHONUM: Reports from the Department of Phonetics, University of Umeå*, 2, 87-106.

Löfqvist, A. (1996). Control of oral closure and release in bilabialstop consonants. *Proceedings of the 6th Australian International Conference on Speech Science and Technology* (pp. 561-566). Canberra.

Löfqvist, A., & Gracco, V. L. (1997). Lip and jaw kinematics in bilabial stop consonant production. *Journal of Speech, Language, and Hearing Research*, 40, 877-893.

Maddieson, I. (2005). Bilabial and labio-dental fricatives in Ewe. *UC Berkeley Phonology Lab Annual Report*, 199-215.

Mermelstein, P. (1973). Articulatory model for the study of speech production. *The Journal of the Acoustical Society of America*, 53, 1970-1082.

Mooshammer, C., Hoole, P., & Geumann, A. (2007). Jaw and order. *Language and Speech*, 50(2), 145-176.

Nam, H. (2007). A competitive, coupled oscillator model of moraic structure: Split-gesture dynamics focusing on positional asymmetry. In J. Cole & J. I. Hualde (Eds.). *Papers in Laboratory Phonology IX* (pp. 483-506). Berlin: Mouton de Gruyter.

Nam, H. (n.d.) Towards articulatory machines. Unpublished manuscript. 남호성. 조음 로봇을 꿈꾸며. 원고.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2253-2256). Barcelona, Spain.

Nam, H., Browman, C., Goldstein, L. Tiede, M., Rubin, P., Proctor, M., Saltzman, E. (2012). TADA (TAsk Dynamics Application) Retrieved from http://www.haskins.yale.edu/tada_download/on September 18, 2017.

Nam, H., Goldstein, L., & Saltzman, E. (in press). Self-organization of syllable structure: a coupled oscillator model. In F. Pellegrino, E. Marisco & I. Chitoran, (Eds.). *Approaches to phonological complexity*. Berlin/New York: Mouton de Gruyter. Retrieved fromhttp://sail.usc.edu/~lgoldste/me/documents/Nam_Gold_Saltz _agent_ES6_LG.pdf on October 20, 2018.

Nam, H., Goldstein, L., Saltzman, E., & Byrd, D. (2004). TADA: An enhanced, portable task dynamics model in MATLAB. *The Journal of the Acoustical Society of America*, 115(5), 2430-2430.

Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In G. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, and Prosody* (pp. 261-280). Cambridge: Cambridge University Press.

Öhman, S. E. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151–168.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., & Jackson, M. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America*, 92, 3078-3096.

R Core Team. (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Retrieved from http://www.rproject.org on November 18, 2017.

Recasens, D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. *The Journal of the Acoustical Society of America*, 76, 1624-1635.

Saltzman, E., & Kelso, J. A. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, 94, 84-106.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1, 333-382.

Saltzman, E., Nam, H., Goldstein, L., & Byrd, D. (2006). The distinctions between state, parameter and graph dynamics in sensorimotor control and coordination. In M. Latash & F. Lestienne, (Eds.). *Motor Control and Learning* (pp. 63-73). New York: Springer.

Shiller, D. M., Houle, G., & Ostry, D. J. (2005). Voluntary control of human jaw stiffness. *Journal of Neurophysiology*, 94, 2207-2217.

Smith, C. (1992). *The timing of vowel and consonant gestures*. Ph.D. Dissertation, Yale University.

Son, M. (2008). Gradient reduction of C1 in /pk/ sequences. *Speech Sciences*, 15(4), 43-65. (손민정 (2008). /pk/연속음에서 일어나는 점층 약화. *음성과학*, 15(4), 43-65.)

Son, M. (2013). Articulatory attributes in Korean nonassimilating contexts. *Phonetics and Speech Sciences*, 5(1), 109-121. (손민정 (2013). 한국어 비동화 환경에서 일어나는 조음 특질. *말소리와 음성과학*, 5(1), 109-121.)

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. (손민정 (2015a). 한국어 /l/ 탄설음화에 나타나는 변이음 [ɾ] 조음 특질: 조음 약화와 조음 중첩. *음성 음운 형태론 연구*, 21, 427-456.)

Son, M. (2015b). Korean /l/-flapping in an /i/-/i/ context. *Phonetics and Speech Sciences*, 7(1), 151-163. (손민정 (2015b). /i/-/i/ 환경에서 일어나는 한국어 /l/ 설탄음화. *말소리와 음성과학*, 7(1), 151-163.)

Son, M. (2018). Word-boundary and rate effects on upper and lower lip movements in the articulation of the bilabial stop /p/ in Korean. *Phonetics and Speech Sciences*, 10(1), 23-31. (손민정 (2018). 한국어 양순 폐쇄음 /p/ 조음에 나타나는 윗입술과 아래입술 움직임에 미치는 단어경계와 속도 효과. *말소리와 음성과학*, 3(4), 33-43.)

Son, M., Kim, S., & Cho, T. (2011). Supralaryngeal articulatory characteristics of coronal consonants /n, t, tʰ, t*/ in Korean. *Phonetics and Speech Sciences*, 3(4), 33-43. (손민정·김사향·조태홍 (2011). 한국어 화관자음 /n, t, tʰ, t*/에서 나타나는 성도 조음 특질. *말소리와 음성과학*, 3(4), 33-43.)

Son, M., Kochetov, A., & Pouplier, M. (2007). The role of gestural overlap in perceptual place assimilation: Evidence from Korean. *Papers in Laboratory Phonology IX* (pp. 507-534). New York: Mouton de Gruyter.

Tiede, M. (2005). MVIEW: Software for visualization and analysis of concurrently recorded movement data. New Haven, CT: Haskins Laboratories.

Wood, S. (1979). A radiographic analysis of constriction location for vowels. *Journal of Phonetics*, 7, 25-43.

• **Son, Minjung**
Professor, Department of English Language and Literature
Hannam University
70 Hannam-ro, Daedoek-gu
Daejeon 34430, Korea
Tel: 042-629-7342
Email: minjungson@hnu.ac.kr
Fields of interest: Phonetics, Phonology