Check for updates

# Word-boundary and rate effects on upper and lower lip movements in the articulation of the bilabial stop /p/ in Korean*

Minjung Son**

## Abstract

In this study, we examined how the upper and lower lips articulate to produce labial /p/. Using electromagnetic midsagittal articulography, we collected flesh-point tracking movement data from eight native speakers of Seoul Korean (five females and three males). Individual articulatory movements in /p/ were examined in terms of minimum vertical upper lip position, maximum vertical lower lip position, and corresponding vertical upper lip position aligned with maximum vertical lower lip position. Using linear mixed-effect models, we tested two factors (word boundary [across-word vs. within-word] and speech rate [comfortable vs. fast]) and their interaction, considering subjects as random effects. The results are summarized as follows. First, maximum lower lip position varied with different word boundaries and speech rates, but no interaction was detected. In particular, maximum lower lip position was lower (e.g., less constricted or more reduced) in fast rate condition and across-word boundary condition. Second, minimum lower lip position, as well as lower lip position, measured at the time of maximum lower lip position only varied with different word boundaries, showing that they were consistently lower in across-word condition. We provide further empirical evidence of lower lip movement sensitive to both different word boundaries (e.g., linguistic factor) and speech rates (e.g., paralinguistic factor); this supports the traditional idea that the lower lip is an actively moving articulator. The sensitivity of upper lip movement is also observed with different word boundaries; this counters the traditional idea that the upper lip is the target area, which presupposes immobility. Taken together, the lip aperture gesture is a good indicator that takes into account upper and lower lip vertical movements, compared to the traditional approach that distinguishes a movable articulator from target place. Respective of different speech rates, the results of the present study patterned with cross-linguistic lenition-related allophonic variation, which is known to be more sensitive to fast rate.

Keywords: upper lip, lower lip, movable articulator, target, lip aperture

## 1. Introduction

From a phonological point of view, place of articulation can be largely classified by the feature [±anterior]: four target regions (i.e., labial, dental, alveolar, and post-alveolar) have the feature [+anterior] feature and five (i.e., palatal, velar, uvular, pharyngeal, epiglottal, glottal) have the feature [-anterior] (Keating, 1991; Clements, 1985; Ladefoged & Maddieson, 1996, *inter alia*). For the

traditional terms listed above, most places of articulation refer to fixed regions along the upper and back surfaces of the vocal tract except for epiglottal (Ladefoged & Maddieson, 1996). At the supralaryngeal level of articulation, constriction occurs by pairing between the place of articulation and a mobile articulator, which is central to generating speech sounds. First, focusing on labial place of articulation, in particular, the upper lip is the target area with two different movable articulators. Labial place is produced with the upper lip as the articulatory objective area involving the active lower lip as a movable articulator (e.g., /p/, /b/). Linguo-labial is produced with the upper lip as the articulatory objective area involving the active tongue blade articulator (e.g., /t̼/, /d̼/). Second, labio-dental is produced with the upper teeth as the articulatory objective area involving the active lower lip (e.g., /ɱ/). However, note that the lip organs are distinguished because both the upper lip and the lower lip are, in principle, physiologically movable regardless of their linguistic status (e.g., an *articulatory target region* with an active articulator involved). Physiologically, lip closing occurs with an elevated (not depressed) mandible and mentalis contraction; earlier in the phonological developmental stage, lower lip raising occurs by mechanically being linked with jaw raising movement. Later in the phonological developmental stage, children, by six years of age, begin to use upper lip lowering in coordination with jaw and lower lip raising movement to make a lip closing gesture beyond passive smacking (Gick *et al*., 2013).

Kinematic characteristics of the lips for bilabial stops in various contexts have been relatively well documented in previous literature. Browman & Goldstein (1988) showed that an onset (e.g., singleton as well as CC(C) clusters) is *globally* organized into a syllable. Examining X-ray microbeam data from Miller & Fujimura (1982), they analyzed vertical movement of the lower lip and the tongue tip, showing that the C-center of C(C(C)) sequences (e.g., 'peak __' (target words: 'pots', 'sots', 'lots', 'spots', 'plots', 'splots') exhibited the most stable temporal relation with an anchor (e.g., the beginning of an acoustic closure for /t/ in the second word) in English. In an electromagnetic articulography study on bilabial /p/ and /b/ in Ewe, Maddieson (2005) examined upper and lower lip movement, where both the upper and lower lips clearly moved faster in voiceless bilabial /p/ than in its voiced counterpart /b/. In terms of gestural spatial magnitude, voiceless bilabial /p/ showed moderately more compression, indicated by lower vertical position of these two articulators on the ordinate plane.

Under the hypothesis of articulatory phonology (Browman & Goldstein, 1986, 1989, 1990, 1992), "gestures are units of action that can be identified by observing the coordinated movements of the vocal tracts" (1989:202). In a computation model (Saltzman *et al*., 1987), a gesture is abstract and invariant at the phonological level of representation and refers to *coordinated movements of articulators* to achieve a linguistically meaningful *task*. Browman & Goldstein (1986) proposed that gestures are understood as tract variables which are specified as constriction location (CL) and constriction degree (CD). Specifically, they are lip protrusion (LP), lip aperture (LA), tongue tip constriction location (TTCL), tongue tip constriction degree (TTCD), tongue body constriction location (TDCL), tongue body constriction degree (TDCD), velic aperture (VEL), and glottal aperture (GLO). The articulators involved in task achievement consist of a set of the upper lip, lower lip, and jaw involved coordinatively in the tract variables of LP and LA, a set of the tongue tip, tongue body, and jaw in the tract variables of TTCL

and TTCD, a set of the tongue body and jaw in the tract variable of TDCL and TDCD, the velum in the tract variable of VEL, and the glottis in the tract variable of GLO. Task-controlled tract variables are assembled, by hypothesis, into a bigger coordinated structure called a *gestural score*. In the *task-dynamics model* of speech production (Saltzman, 1986; Saltzman & Kelso, 1987), each gesture is mathematically represented using several parameters (e.g., mass, damping, and stiffness) and temporal periods are also specified for a given utterance (see Saltzman & Kelso (1987) for parameters of a set of equations and Nam (manuscript) for a review of task-dynamics). Articulatory studies have been done to serve the purpose of providing kinematic data entered for parameter values entered in gestures and gestural scores (Browman & Goldstein, 1995, 1988).

Lip gestures have been well studied in various languages using various methodologies tracking articulatory movements (see Browman & Goldstein's (1986, 1988, 1990, 1995) x-ray microbeam study on English and Chaga; Kochetov *et al*.'s (2007) electromagnetic articulography study on Korean and Russian; Löfqvist's (1996) and Löfqvist & Gracco's (1997) simultaneous electromagnetic articulography and aerodynamic study on American English; Maddieson's (2005) electromagnetic articulography study on Ewe; Smith's (1992) x-ray microbeam study on Japanese and Italian; Ladefoged & Maddieson's (1996) videotape study on Vao; Son's (2008, 2013), Son *et al*.'s (2007), and Son *et al*.'s (2012) electromagnetic articulagrphy study on Korean; Yanagawa's (2006) electromagnetic articulography study on American English, Cantonese, Taiwanese, German, French, and Japanese). For the analysis of kinematic data, the lip aperture gesture was analyzed for Korean. Using an electromagnetic midsagittal articulometer study, Son (2008) observed spatio-temporal reduction in the LA gesture in assimilating context such as /Vp(#)kV/ sequences with inter-speaker variability, indicating that more reduction of the LA gesture was observed in fast and within-word boundary conditions if it ever occurred. Comparing Korean three-way laryngeal contrast (lenis, fortis, aspirated) in four vocalic contexts (/iCi/, /aCa/, /iCa/, /aCi/), Son *et al*. (2012) examined kinematic data from a set of nonsense words with bilabial stops. Analyzing the LA gesture, they showed that aspirated stop /pʰ/ demonstrated a greater lip closing displacement than lenis stop /p/ (/pʰ/>/p/). In lip opening movements, fortis /p*/ was always greater in spatial displacement, acceleration duration, and overall movement duration (/p*/>/p/). Comparing the high vowel context and low vowel context, they consistently observed greater spatial displacement as well as peak velocity in both lip closing and opening movements (/aCa/>/iCi/). Browman & Goldstein (1995) also used lip aperture in the analysis of x-ray microbeam data from one subject. They examined lip constriction, which was estimated along with its corresponding LA gesture value in the analysis of syllable position effects. Systematically varying the location of pitch accents in either one of words in a carrier phrase (e.g., '**MY** __ huddles/puddles/tuddles'; 'my __ **HUDDLES/PUDDLES/TUDDLES**') or in a target word ('**POP**', '**TOT**', '**CAULK**'), bilabial stop /p/ was more spatially reduced in coda position, coherently patterning together with coronal and velar. In addition, they sometimes used vertical lower lip movements for the sake of graphic compatibility with other vertical movements with the tongue tip or tongue body when it was not the focus of analysis (e.g., 'say *leap* again' for onset /l/ vs. 'say *peel* again' for coda /l/).

The individual articulatory movements of the two lips were also the main focus of analysis in a simultaneous electromagnetic articulography and air pressure study. Löfqvist (1996) analyzed the upper lip, lower lip, and jaw for voiced/voiceless bilabial stops (/b/, /p/) elicited within a carrier phrase (e.g., 'say __ again') with reference to acoustic wave forms and air pressure data. In his study, the lip aperture kept changing during an acoustic silence as he analyzed kinematic data from four subjects (three American native speakers and one Swedish native speaker). To quote Löfqvist (1996:563), "The receivers continue to move during the closure due to their placement and to compression of the lip tissues...the lips may be meeting at a high velocity and also that there may be a mechanical interaction between the lips during the closure." In particular, the minimum vertical position of the upper lip coincided with neither the lip aperture minima nor the maximum vertical position of the lower lip, since the upper lip reached the lowest position at the time of target attainment and then receded upward as it yielded to the ongoing lower lip raising movement until the latter reaches its maximum vertical position. This was further related to negative lip aperture. As noted in Gick *et al*. (2013), overshoot (e.g., negative lip aperture) occurs in any constriction since speakers can achieve constriction between a moving articulator and its target location without delivering obligatory *fine control* by speakers.

It seems to be sufficient if articulatory studies provide empirical data for determining parameter values for gestures and gestural scores in the *task-dynamics model* of speech production (Saltzman, 1986; Saltzman & Kelso, 1987). In this regard, the LA gesture in Korean has been relatively well documented in three-way laryngeal contrast and place assimilation. However, it is yet to be discovered, as was done with American English (Löfqvist, 1996), how the upper and lower lips articulate to produce labial consonants, which will ultimately enhance our understanding in a more comprehensive way. In this study, we proceed with this line of research by determining what happens at the time of maximum contact of the upper and lower lips.

### 1.1. Research questions

The production of an utterance is hypothesized to be processed along a succession of different linguistic stages (i.e., lexicon - morphology - syntax - phonology - phonetic execution in the vocal tract) (Mihalicek & Wilson, 2011). In transformational generative grammar, an utterance is decomposed into a syntactic structure (Radford, 1988). Syntactically, the subject occurs in the specifier of an inflectional phrase (IP), which in turn precedes the complement of the IP. Using X-bar schema (Chomsky, 1993), the subject is followed by the object (e.g., [IP [NP [N' [N na]]] [I' [VP [V' [NP [N' [N pap]]] [V məknɨnta]]]]] 'I eat (a bowl of) rice') in a head-final language like Korean. Structurally, the subject and the object are governed by distinct maximal projections, two NPs, and a word boundary occurs between these two NPs.

For Korean, the spatio-temporal reduction of the lip aperture gesture in terms of LA minima was attributed to place assimilation in /...apka.../ sequences, along with more gestural overlap with inter-speaker variability (Son *et al*., 2007). What they observed from an electromagnetic midsagittal articulometer study was categorically reduced LA in within-word condition, which always occurred within an accentual phrase, although not all /pk/ clusters within an accentual phrase demonstrated reduced LA gestures (see Jun (1993, 2006) for intonational prosodic structure of Seoul Korean). In

contrast, Jun (1996) observed partially or fully reduced lip gestures even in across-word boundary condition from an aerodynamic study on place assimilation in Korean. In another electromagnetic articulography study (Son, 2008), the target of place assimilation (e.g., /...Vp(#)kV.../) generally showed less constriction in lenis bilabial stop /p/ in the within-word condition, compared to across-word boundary, but there was no partially or fully reduced LA for most speakers (four out of five speakers indicated less constriction in the within-word condition). Meanwhile, one exception was found in one speaker out of five speakers, who showed more reduction in across-word condition and exhibited partially or fully reduced LA in that context if there was any reduction at all.

Korean also showed more spatio-temporal gestural reduction in coda compared to onset in an electromagnetic midsagittal articulometer study with regard to constriction duration as well as constriction degree (e.g., syllable-initial /k/ > syllable-final /k/ in Son (2011)). This is compatible with the results in Browman and Goldstein's (1995) X-ray microbeam study of American English where more gestural reduction in coda was observed within a single identical prosodic domain (e.g., syllable-initial /p/, /t/, /k/ > syllable-final /p/, /t/, /k/).

In this paper, we examine word boundary effects (e.g., linguistic factor) on gestural reduction confined to intervocalic syllable-initial onset as we focus on Korean bilabial lenis stop /p/ in two morpho-syntactic contexts (within-word boundary vs. across-word boundary). To serve this purpose, we proceed with lenis bilabial stop /p/ flanked by a set of homorganic low vowels (/...a(#)Ca.../), where C occurs consistently in syllable-initial position in the Korean orthography.

A further goal is to learn whether rate-dependent articulatory movement occurs in the upper lip and/or lower lip movements. Son's (2008) electromagnetic articulography study showed that the lip aperture minima indicated less constriction for labial place occurring in assimilating context /...ap(#)ka.../ in fast speech rate and with inter-speaker variability. Meanwhile, an intervocalic lateral /l/, which is phonetically executed as a flap /ɾ/ in Korean, did not demonstrate such speech rate effects on constriction degree when it was evaluated with vertical tongue tip position, not only in the low vowel context (/...ala.../) but also in the high vowel context (/...ili.../) (Son, 2015a, 2015b). In an effort to enhance our understanding of speech rate effects (e.g., paralinguistic factor) on constriction degree, on the one hand, and individual articulators involved in constriction degree, on the other, we pursue an examination of minimum vertical position of the upper lip (i.e., lowering movements) and the maximum vertical position of the lower lip (i.e., raising movements), while factoring in two speech rates (comfortable vs. fast).

Taken together, we examine maximum vertical position values of the lower lip and two minimum values of the lower lip (minimum vertical position values of the lower lip and corresponding vertical position values lined up with the maximum vertical position of the lower lip) in two word boundary conditions (across-word vs. within-word) and two speech rate conditions (comfortable vs. fast).

## 2. Method

### 2.1. Participants

Eight (three male and five female) native Seoul-Korean speakers voluntarily participated in the electromagnetic articulometer study (i.e., flesh-point tracking system) and were financially rewarded[1]. At the time of data collection, the subjects were in their mid-twenties and early thirties, engaged in their graduate studies, and had spent their first twenty years of life in Seoul or Gyeonggi province in South Korea. Also at the time of data collection, they all resided in Connecticut, U.S.A., and were not isolated from Korean communities. They all identified themselves as native speakers of Seoul Korean, without any temporary or permanent speech or hearing impairment in the past.

### 2.2. Data collection and stimuli

Electromagnetic midsagittal articulometer (EMMA in Perkell *et al.*, 1992) was used to derive kinematic data relating to articulator movement. The two-dimensional point-tracking system records the positional values of electric transducers (i.e., receiver coils) attached to several articulators: the upper lip, the lower lip, the tongue tip, the tongue body, the tongue dorsum, and the lower incisor. Three transmitters secured on a plastic helmet generate a magnetic field alternating at different frequencies and induce an alternating current in the transducers. This enables us to extract the distances of each of the transducers from the three transmitters, which are in turn expressed as a vector on an ordinate plane (see also Löfqvist (1993), for a more detailed description of how the electromagenetic transduction technique works). Articulatory data was sampled at 200 Hz (i.e., one frame collected every 5 milliseconds) and further smoothed by a low-pass filter of 20 Hz using post-processing procedures in Matlab software by Mathworks. For the purpose of the current study to examine the upper and lower lip movements, we limited the scope of analysis to the kinematic characteristics of the vertical movement of the two lips. Acoustic data was also acquired simultaneously at the time of articulatory data collection.

(1) Stimuli
a. Target sequence /pa/

   i. Within-word boundary condition.
   /a**p**ai̇/ 'father' (North Korean dialect)

   ii. Across-word boundary condition.
   /**p**akat∫i̇/ '(a) gourd dipper'

b. Natural short sentence including the target sequence and its syntactic structure using maximal projections (following Chomsky (1993) and simplified) shown in (a.i) and (a.ii). The symbol '#' represents a word boundary.

   i. Within-word boundary condition.
   /a**p**ai̇ # toŋmuni̇n # pukhanmalija/
   [IP[NP a**p**ai̇ toŋmuni̇n]   [VP[NP pukhanmal]   [V ija]]]

*'Father comrade* is North Korean vocabulary.'

   ii. Across-word boundary condition.
   /t∫əna # **p**akat∫i̇li̇ # pala/
   [IP[NP t∫əna]   [VP[NP **p**akat∫i̇li̇]   [V pʰala]]]
   'Jeona sells gourd dippers.'

We presented stimuli containing the target sequence to subjects as the presentation of stimuli was blocked by word boundary (syntactic structures are used to create different word boundaries (across-word boundary vs. within-word boundary)) and speech rate conditions. The across-word boundary condition was always acquired before the within-word boundary condition, and comfortable speech rate before fast speech rate. Eight subjects were instructed to read a given short natural sentence eight times, with the first four repetitions interrupted by four repetitions of a different short natural sentence; in total, eight repetitions for a given target word or phrase were acquired for further analysis. A total of 223 tokens from seven speakers were available for further analysis, since data from one female subject and one token from one male speaker were not included due to various reasons (e.g., stuttering and a data conversion problem). A set of stimuli within a presentation block was randomly ordered, which was consistent across blocks as well as subjects. Both stimuli sequences occurred in a potentially weakening environment, i.e., consistently in the intervocalic position.

### 2.3. Measurements

Using the function of *lp_Snapex* in MVIEW (Tiede, 2005), we demarcated minimum vertical upper lip position and maximum vertical lower lip position relevant to the articulation of bilabial stop /p/. In particular, *lp_Snapex* used an algorithm based on velocity profiles and determined zero velocity. Also determined was corresponding vertical upper lip position as lined up with maximum vertical lower lip position. <Figures 1.a.i, 1.b.ii, and 1.c.iii> illustrate the respective specifics of the gestural demarcation superimposed on identical realtime movement trajectories of the upper and lower lips. Each figure shows one selected window of an across-word boundary /a # pa/ sequence and is captured from the temporal display in MVIEW.

### 2.4. Statistical analysis

As we take into account individual participant differences, we conducted linear mixed effects models in R (R Development Core Team, 2014). The results of articulatory analysis were fitted with the lmer function from the lme4 package (Bates *et al.*, 2011). For analysis, we took as dependent variables the minimum vertical upper lip position <Figure 1.a.i>, maximum vertical lower lip position <Figure 1.b.ii>, and corresponding vertical upper lip position lined up with maximum vertical lower lip position <Figure c.iii>. We used the word boundary (across-word boundary vs. within-word boundary) and speech rate (comfortable vs. fast) conditions as fixed factors, and subjects (7 subjects) as a random factor. We draw a comparison between results of null models from mixed effects analysis and that of a full model (Speech rate +

Boundary) in evaluating main effects, as well as between the result of a reduced model (Speech rate X Boundary) and that of a full model (Speech rate + Boundary) in evaluating an interaction.
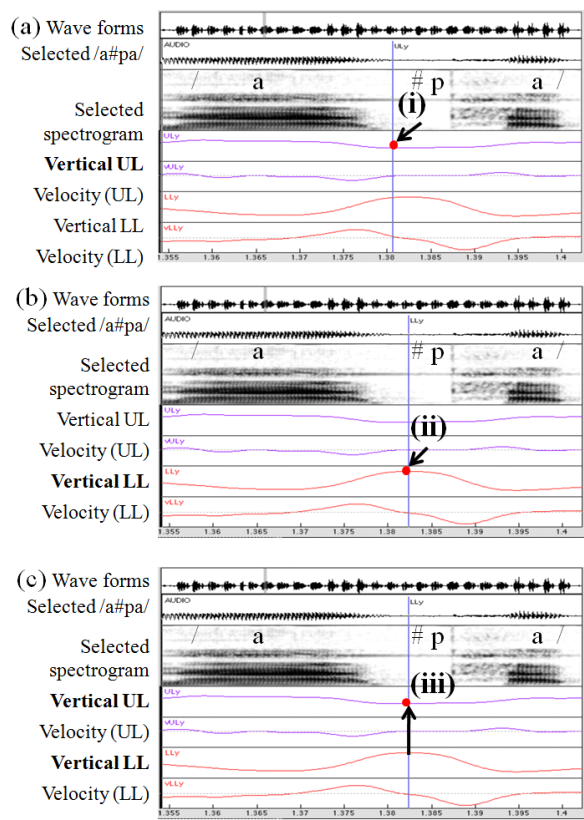


**Figure 1.** Greater values denote higher position for the upper and lower lip movements. (a) specifies minimum vertical upper lip in (i). (b) specifies maximum vertical upper lip in (ii). (c) specifies corresponding vertical lower lip position lined up with the vertical maxima of the lower lip movement in (iii). The captured window depicts the first token of the across-word boundary condition at comfortable rate produced by a female speaker (SF1).

## 3. Results

### 3.1. Vertical lower lip maxima

Results showed no interaction between Speech rate and Boundary ($\chi^2$=1.69, $p$>0.05). Adding Speech rate <Table 1.c.i>, we shifted the variance that was previously seen in the random effects in the null model as shown in <Table 1.b.i> (the residual of 2.45 stands for the random variation that does not stem from the individual subjects component) to the fixed effects component (the residual of 2.37). To conclude, the vertical lower lip maxima varied with different speech rates ($\chi^2$= 6.73, $p$<0.01), lowering it by 0.54 mm ($SE$. ±0.21) at fast rate (comfortable>fast).

In terms of different word boundaries, vertical lower lip maxima also varied with Boundary ($\chi^2$=7.36, $p$<0.01). Including Boundary <Table 1.c.i>, we shifted the variance that was previously seen in the random effects <Table 1.a.i> (the residual of 2.46 stands for the random variation that does not stem from the individual subjects component) to the fixed effects component (the residual of 2.37). To conclude, speakers raise the lower lip at maximum constriction by 0.57 mm ($SE$. ±0.21) in the within-word condition (across-word < within-word). In the following, the result of Speech rate is shown in <Figure 2.a> and that of Boundary in <Figure 2.b> using a box plot.

**Table 1.** Results of linear mixed effects models
*Note*: Number of observations: 223. Groups: subject, 7

**Null model**
**a.i.** Random effects: Null model (Speech rate)

| Groups | Name | Variance | *SD* |
|---|---|---|---|
| Subject | (Intercept) | 0.97 | 0.98 |
| Residual | | 2.46 | 1.57 |

**a.ii.** Fixed effects: Null model (Speech rate)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -15.33 | 0.40 | -38.48 |
| Speech rate [fast] | -0.54 | 0.21 | -2.56 |

**b.i.** Random effects: Null model (Boundary)

| Groups | Name | Variance | *SD* |
|---|---|---|---|
| Subject | (Intercept) | 0.96 | 0.98 |
| Residual | | 2.45 | 1.57 |

**b.ii.** Fixed effects: Null model (Boundary)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -15.88 | 0.40 | -39.84 |
| Boundary [within-word] | 0.56 | 0.21 | 2.68 |

**Full model**
**c.i.** Random effects: Full model (Speech rate + Boundary)

| Groups | Name | Variance | *SD* |
|---|---|---|---|
| Subject | (Intercept) | 0.96 | 0.98 |
| Residual | | 2.37 | 1.54 |

**c.ii.** Fixed effects: Full model (Speech rate + Boundary)

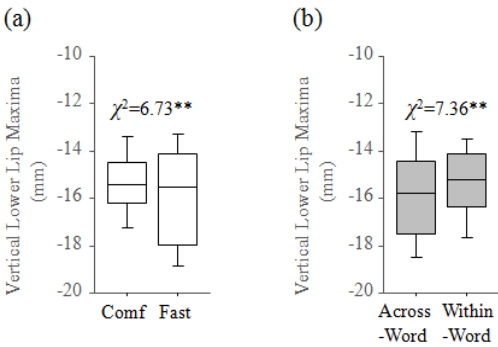| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -15.61 | 0.41 | -38.00 |
| Speech rate [fast] | -0.54 | 0.21 | **-2.62** |
| Boundary [within-word] | 0.57 | 0.21 | **2.74** |



**Figure 2.** Vertical lower lip maxima (a) Speech rates and (b) Boundary types. (The symbol '**' is for $p$<0.01.)

### 3.2. Vertical upper lip minima

There was no interaction between Speech rate and Boundary ($\chi^2$= 0.18, $p$>0.05). Adding Boundary <Table 2.c.i>, we shifted the variance that was previously seen in the random effects <Table 2.a.i> (the residual of 2.94 stands for the random variation that does not stem from the individual subjects component) to the fixed effects component (the residual of 2.84). To conclude, speakers lower the

upper lip at minimum constriction by 0.63 mm (*SE*. ±0.23) in the across-word condition (across-word < within-word) ($\chi^2$=7.75, *p*<0.01). However, vertical upper lip minima were not influenced by Speech rate ($\chi^2$=0.74, *p*>0.05). In <Figure 3>, the results of Speech rate and Boundary are separately plotted using a box plot.

**Table 2.** Results of linear mixed effects models
*Note*: Number of observations: 223. Groups: subject, 7

**Null model**
**a.i.** Random effects: Null model (Speech rate)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.83 | 2.80 |
| Residual | | 2.94 | 1.72 |

**a.ii.** Fixed effects: Null model (Speech rate)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -0.01 | 1.07 | -0.01 |
| Speech rate [fast] | -0.19 | 0.23 | -0.83 |

**b.i.** Random effects: Null model (Boundary)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.83 | 2.80 |
| Residual | | 2.85 | 1.69 |

**b.ii.** Fixed effects: Null model (Boundary)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -0.42 | 1.07 | -0.40 |
| Boundary [within-word] | 0.63 | 2.23 | 2.80 |

**Full model**
**c.i.** Random effects: Full model (Speech rate + Boundary)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.83 | 2.80 |
| Residual | | 2.84 | 1.69 |

**c.ii.** Fixed effects: Full model (Speech rate + Boundary)

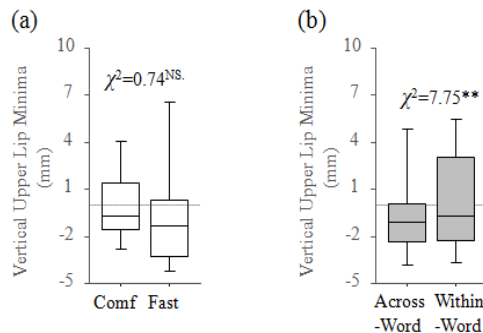| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -0.33 | 1.08 | -0.30 |
| Speech rate [fast] | -0.19 | 0.23 | -0.86 |
| Boundary [within-word] | 0.63 | 0.23 | **2.81** |



**Figure 3.** Vertical upper lip minima (a) Speech rates and (b) Boundary types. (The symbol '**' is for *p*<0.01.)

### 3.3. Corresponding vertical upper lip position lined up with vertical lower lip maxima

In evaluating the dependent variable of corresponding vertical upper lip position measured by aligning it with vertical lower lip maxima, neither interaction between Speech rate and Boundary nor significant effects of Speech rate were observed ($\chi^2$=0.02; $\chi^2$=1.12, all at *p*>0.05). Including Boundary <Table 3.c.i>, we shifted the variance that was previously seen in the random effects <Table 3.a.i> (the residual of 2.87 stands for the random variation that does not stem from the individual subjects component) to the fixed effects component (the residual of 2.77). To conclude, speakers lower the upper lip by 0.61 mm (*SE*. ±0.22) in the across-word condition (across-word < within-word) ($\chi^2$=7.27, *p*<0.01). In <Figure 4>, the results of Speech rate and Boundary type are separately plotted using a box plot.

**Table 3.** Results of linear mixed effects models
*Note*: Number of observations: 223. Groups: Subject, 7

**Null model**
**a.i.** Random effects: Null model (Speech rate)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.79 | 2.79 |
| Residual | | 2.87 | 1.69 |

**a.ii.** Fixed effects: Null model (Speech rate)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | 0.13 | 1.07 | 0.12 |
| Speech rate [fast] | -0.23 | 0.23 | -1.03 |

**b.i.** Random effects: Null model (Boundary)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.79 | 2.79 |
| Residual | | 2.79 | 1.67 |

**b.ii.** Fixed effects: Null model (Boundary)

| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -0.29 | 1.07 | -0.27 |
| Boundary [within-word] | 0.61 | 0.22 | 2.71 |

**Full model**
**c.i.** Random effects: Full model (Speech rate + Boundary)

| Groups | Name | Variance | SD |
|---|---|---|---|
| Subject | (Intercept) | 7.79 | 2.79 |
| Residual | | 2.77 | 1.67 |

**c.ii.** Fixed effects: Full model (Speech rate + Boundary)

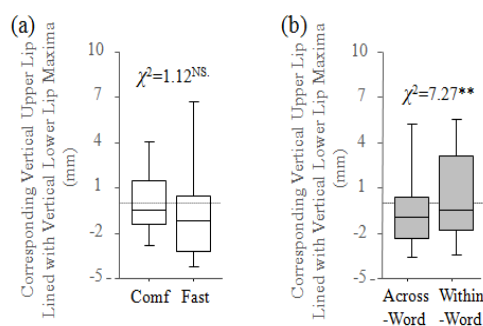| | Estimate | *SE* | *t*-value |
|---|---|---|---|
| (Intercept) | -0.18 | 1.07 | -0.16 |
| Speech rate [fast] | -0.24 | 0.22 | -1.06 |
| Boundary [within-word] | 0.61 | 0.22 | **2.72** |

**Figure 4.** Corresponding vertical upper lip lined with vertical lower lip maxima (a) Speech rates and (b) Boundary types. (The symbol '**' is for *p*<0.01.)

## 4. Summary and discussion

Vertical lower lip maxima varied with different word boundaries, manifesting more reduction of the lower lip movement in the across-word condition. In contrast, the upper lip moved to a greater extent downwards in the across-word boundary condition in terms of minimum vertical upper lip position and corresponding vertical upper lip position measured at the time point of lower lip maxima; an articulatorily adaptive compensation might have occurred as a concurrent reaction to articulatory reduction of the lower lip. In addition, more reduction in the fast rate condition was exhibited, being confined to maximum vertical lower lip position; this is compatible with previous studies on lenition-related allophonic variations (Kirchner, 1998), complying with relatively more or frequent gestural weakening in fast rate.

### 4.1. Lip aperture as the subject of articulatory study

A bilabial stop takes the form of a tight seal between the pairing of the movable upper and lower lips, being concurrent with an overshooting (i.e., negative lip aperture) and a spread constriction of the lips (Ladefoged & Maddieson, 1996; Gick *et al*., 2013, *inter alia*). In traditional accounts, when articulating bilabial stops, the upper lip is the target place of articulation toward which the lower lip actively changes position (Ladefoged & Maddieson, 1996). Testing word boundary effects as well as speech rate effects in vertical lower lip maxima, the results revealed that the lower lip is a movable articulator, varying with a linguistic factor (across-word boundary < within-word boundary) and a paralinguistic factor (comfortable > fast) that we were interested in. Contrary to traditional accounts, the upper lip, known as the target location of constriction in labial place, also varied with different word boundaries; therefore, the veracity of the upper lip being the target region has not been confirmed. Taken together, our current study of the bilabial voiceless stop /p/ provided empirical evidence to uphold the assertion of articulatory phonology that articulatory characteristics should be described by reference to the lip aperture (LA) gesture which includes the upper lip and lower lip (as well as the jaw) as articulators (Browman & Goldstein, 1986, 1989, 1990, 1992).

With respect to moving articulators, it is intuitive to assume greater spatial displacement of the lower lip during lip closing and opening movement since it has the specified shape of the mandibular prominence. The mandible is of use in articulating vowels and consonants (Wood, 1979; Satzman & Munhall, 1989; Browman & Goldstein, 1990; Mooshammer *et al*., 2007). The jaw

height is known to differ depending on constriction degree of vowels (Wood, 1979) and consonantal types with different manner of articulation (e.g., gradual increasing in the order /t/ > /d/, /n/, /l/ in Keating *et al*. (1994) and Mooshammer *et al*. (2003); no change among /t/, /d/, /s/, and /ʃ/, but gradual increasing in the order loud speech > comfortable speech, being confined to /n/, and higher in the order /t/, /d/, /s/, /ʃ/ > /l/ with inter-speaker variability in Mooshammer *et al*. (2007)). However, examining a set of coronal stop consonants, Son *et al*. (2011) found invariable vertical jaw maxima among coronal /t/, /t*/, /tʰ/, and /n/ in homorganic low vowel context with nonsense words (/aCa/). Likewise, speech rate effects were absent from constriction maxima in vertical tongue tip gesture in a high vowel context as well as low vowel context (fast=comfortable in /ala/→[aɾa] in Son (2015a); fast=comfortable in /ili/→[iɾi] in Son (2015b)). In sum, there existed some segment-specific or speech style-dependent jaw height difference despite the absence of coherent results across studies. To add to our understanding of how the jaw articulator behaves in terms of configuring a segment in Korean, it is of use to resolve related questions about jaw articulator movement in future study, as follows; i) does it serve consonantal articulation limited to functional movement of an active articulator that is elevated upwards to form constriction? (Satzman & Munhall, 1989; Browman & Goldstein, 1990), and ii) does it do so possibly beyond a simple assistance by manifesting diverse jaw position with different manner of articulation and/or jaw height difference of a single segment with different linguistic contexts (e.g., syntax, prosodic structure, speech rate/style) (Mooshammer *et al*., 2007)? Since these questions are beyond the scope of current study, we leave them for further analysis.

### 4.2. Articulatory reduction of the lower lip in the across-word boundary condition

We did observe that the vertical lower lip maxima was lower (more reduction) in the across-word boundary condition, compared to the within-word condition. Comparable results were obtained in some previous studies. Byrd's (1996) electropalatography study on intergestural coordination showed, confined to one speaker (out of five speakers) in its occurrence, observed a trend of less constriction in onset /k/ if a /Vs#kV/ sequence is broken up by an immediately preceding word boundary, compared to /V#skV/ sequences and /Vsk#V/ sequences (e.g., '*Type a scrab again.*', '*Type bass cap again.*', '*Type mask amp again.*'), but the opposite pattern was not observed across speakers. We suggest that the twofold functions of oral constriction gestures (e.g., consonantal and vocalic gestures in Browman & Goldstein (1992)) may be, in part, attributed to more gestural reduction of intervocalic onset in the across-word condition. In a *task-dynamics model* of speech production, the timing of a gesture is governed by coupling oscillators (Saltzman & Munhall, 1989). By hypothesis, coordination among gestures can exhibit stable modes (in-phase, 0°) (to which CV sequences belong), which can be represented using coupling graphs (Goldstein *et al*., 2006). Since the target /p/ from the current study occurs in onset position irrespective of different word boundary conditions, this implies that a within-word /apa/ sequence is not distinguished from an across-word /a#pa/ sequence in terms of coupling graphs; constriction actions and their relative phasing between /p/ and /a/ are both invariant. However, given that V-to-V coarticulation occurs due to the twofold functional characteristics of oral gestures - consonantal and vocalic (Browman & Goldstein, 1992; Öhman, 1967), we conjecture that

two instances of /a/ interrupted by a word boundary may have caused less stable modes in term of V-to-V coarticulation by gestural blending (Romero, 1996), in part, which may have in turn induced gestural spatial reduction such as that observed in the current study. We may find a more hint about more reduction in the across-word boundary condition to the general assumption of articulatory phonology; intergestural coordination mode is specified in lexical items and extensively applied to a sequence of lexical items, but articulatory consequence can vary (Browman & Goldstein, manuscript). We will leave this issue for further study.

In addition, caution should be taken since the result of the current study of the bilabial voiceless stop /p/ revealed that the lip aperture gesture, as a holistic measure, is more appropriate as the subject of articulatory study. In future study, linguistic and paralinguistic factors need to be evaluated in terms of the lip aperture gesture in intervocalic position, which ultimately conforms to the assertion of articulatory phonology, "gestures are the units of action that can be identified by observing the coordinated movements of the vocal tracts" (Browman & Goldstein, 1989:202).

We would like to conclude this study by pointing out a deficiency that needs to be improved upon in future study. Note that one set of target onsets from our data did not exhibit *paradigmatic* contrast in terms of syntactic or prosodic hierarchical structures; the target onset in the within-word condition is part of a word in sentence-initial or utterance-initial position (e.g., [$_{\text{InflP(IP)}}$[$_{\text{NP(AP)}}$ σσσ σσσ ...), while the target onset in the across-word condition is part of a word which is one word away from the sentence-initial or utterance-initial position ([$_{\text{InflP(IP)}}$[$_{\text{NP(AP)}}$ σσ [$_{\text{NP(AP)}}$ **σ**σσσσ ... or [$_{\text{InflP(IP)}}$[$_{\text{NP(AP)}}$ σσ [$_{\text{NP}}$ **σ**σσσσ ...). The asymmetric distribution of stimuli stems from the fact that we have segmental contexts and syllabic structure balanced such that a lenis stop target /p/ occurs in intervocalic position (e.g., /a(#)Ca/ as a possible lenition context) and occupies onset position orthographically. Nevertheless, admitting that the stimuli conditioning was not prepared especially for us to systematically evaluate word boundary effects on articulatory reduction, neither did we find apparent evidence that word-internal onset is characterized by a strengthening/lengthening position compared to word-initial onset, one word apart from sentence-/utterance-initial position (cf., domain-initial strengthening in the order Ui>IPi>APi>Wi in Cho & Keating (2001); domain-edge lengthening in the first segment of an accentual phrase and in the final syllable of an intonational phrase in Jun (1993)).

## Acknowledgements

## References

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

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

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

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

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

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

Browman, C., & 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., & Goldstein, L. *Articulatory phonology*. Unpublished manuscript.

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

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

Chomsky, N. (1993). *Lectures on government and binding: The Pisa lectures*. (No. 9). Dordrecht: Foris.

Clements, G. (1985). The geometry of phonological features. *Phonology Yearbook*, 2, 225-252.

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

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215-249). Cambridge: Cambridge University Press.

Jun, J. (1996). Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology*, 13, 377-407.

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

Jun, S-A. (2006). Korean intonational phonology and prosodic transcription. In S-A. Jun (Ed.) *Prosodic typology: The phonology of intonation and phrasing* (pp. 201-229). Oxford: Oxford University Press.

Keating, P. (1991). Coronal places of articulation. In C. Paradis, & J.-F. Prunet (Ed.), *The special status of coronals: Internal and external evidence* (pp. 29-48). San Diego: Academic 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.

Kirchner, R. (1998). *An effort-based approach to consonant lenition*. Ph.D. Dissertation. University of California in Los Angeles.

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

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 bilabial stop consonants. *Proceedings of the 6th Australian International*

*Conference on Speech Science and Technology* (pp. 561-566). Canberra.

Löfqvist, A., & Gracco, V. (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.

Mihalicek, V., & Wilson, C. (2011). *Language files: Materials for an introduction to language and linguistics*. Columbus: Ohio State University Press.

Miller, J., & Fujimura, O. (1982). Graphic displays of combined presentations of acoustic and articulatory information. *Bell Labs Technical Journal*, 61, 799-810.

Mooshammer, C., Geumann, A., Hoole, P., Alfonso, P., van Lieshout, P., & Fuchs, S. (2003). Coordination of lingual and mandibular gestures for different manners of articulation. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 81-84). Barcelona.

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

Öhman, S. (1967). Numerical model of coarticulation. *The Journal of the Acoustical Society of America*, 41(2), 310-320.

Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., & Jackson, M. (1992). Electromagnetic midsagittal articulometer (EMMA) 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. Available at http://www.rproject.org.

Radford, A. (1988). *A transformational grammar: A first course*. Cambridge: Cambridge University Press.

Romero, J. (1996). Articulatory blending of lingual gestures. *Journal of Phonetics*, 24, 99-111.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. In H. Heuer & C. Fromm (Eds.), *Generation and Modulation of Action Patterns* (pp. 129-144). New York: Springer.

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

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

Saltzman, E., Rubin, P., Goldstein, L., & Browman, C. (1987). Task-dynamic modeling of intergestural coordination (Abstract). *The Journal of the Acoustical Society of America*, 82 (Suppl. 1), 515.

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

Son, M. (2008). Gradient reduction of $C_1$ in /pk/ sequences. *Speech Sciences*, 15(4), 43-65. (손민정 (2008). /pk/연속음에서 일어나는 점층 약화. *음성과학*, 15(4), 43-65.)

Son, M. (2011). Coordinations of articulators in Korean place assimilation. *Phonetics and Speech Sciences*, 3(2), 29-35. (손민정 (2011). 한국어 위치 동화에서 일어나는 조음 협력. *말소리와 음성과학*, 3(2), 29-35.)

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, 427-456. (손민정 (2015a). 한국어 /l/ 탄설음화에 나타나는 변이음 [ɾ] 조음 특질: 조음 약화와 조음중첩. *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., 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., 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., Pouplier, M., & Kochetov, A. (2007). The role of gestural overlap in perceptual place assimilation: Evidence from Korean. *Papers in 9th Conference on 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.

Yanagawa, M. (2006). *Articulatory timing in first and second language: A cross-linguistic study*. Ph.D. Dissertation, Yale University.

남호성. *Towards articulatory machines*. Unpublished manuscript. *조음 로봇을 꿈꾸며*. 원고.

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