



Open Library of Humanities

## Preboundary lengthening and articulatory strengthening in Korean as an edge-prominence language

**Jonny Jungyun Kim**, Department of English Language and Literature, Pusan National University, Busan, Korea, [jonnykkim@gmail.com](mailto:jonnykkim@gmail.com)

**Sahyang Kim**, Department of English Education, Hongik University, Seoul, Korea, [sahyang@hongik.ac.kr](mailto:sahyang@hongik.ac.kr)

**Taehong Cho\***, Department of English Language and Literature, Hanyang University, Seoul, Korea; Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, Korea, [tcho@hanyang.ac.kr](mailto:tcho@hanyang.ac.kr)

\*Corresponding author.

---

Preboundary lengthening, a phenomenon observed in various languages, is examined here in Korean, an ‘edge-prominence’ language that uses phrasing to indicate prosodic groupings and prominence. It was hypothesized that preboundary lengthening in Korean would involve hyperarticulation, reflecting prominence in both spatial and temporal dimensions at the right edge of a phrase. Our measurements of preboundary lengthening, using labial closing and opening durations in both CVCV# and CVCVC# (where ‘#’ = a prosodic boundary), revealed its leftward extension to the initial syllable in bisyllabic words, with the greatest temporal magnitude observed on the final syllable. Crucially, preboundary lengthening was found to be accompanied by a substantial increase in displacement and peak velocity, especially in gestures related to the final syllable. This observation indicates domain-final articulatory strengthening, a phenomenon, largely dovetailed with the notion of an edge-prominence language, and not typically observed in ‘head-prominence’ languages (e.g., English), where the stressed syllable serves as the locus of hyperarticulation. These results were compared in two different prominence conditions driven by information structure (‘new’ vs. ‘given’) and were discussed to understand the observed kinematic pattern in dynamical terms in the theoretical framework of the  $\pi$ -gesture model.

---



## 1. Introduction

Preboundary lengthening (PBL) is a general cross-linguistic propensity of articulatory slowing down toward the end of a phrase (Lindblom, 1968; Byrd & Saltzman, 2003). A growing body of studies on PBL, however, have demonstrated that its phonetic implementation is fine-tuned in a language-specific way, revealing variation and universals in PBL (e.g., Edwards, Beckman, & Fletcher, 1991; Gussenhoven & Rietveld, 1992; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992; Berkovits, 1993; Berkovits, 1994; Byrd, 2000; Cambier-Langeveld, 2000; Byrd, Krivokapić, & Lee, 2006; Cho, 2006; Turk & Shattuck-Hufnagel, 2007; Katsika, 2016; Seo, Kim, Kubozono, & Cho, 2019; see also Fletcher, 2010; Cho, 2015; Cho, 2016 for a review). A great deal of previous studies on PBL, however, have been based on acoustic data in a relatively small number of languages. There is, therefore, much left to understand regarding its articulatory-kinematic underpinnings within and across languages. In fact, due to a limited access to the apparatus for articulatory data collection, there is not enough articulatory data available to understand speech production in general from articulatory perspectives. The present study builds on the PBL literature by providing articulatory-kinematic data associated with PBL in Seoul Korean (henceforth Korean), and by examining how PBL interacts with the prominence system of the language with a view of understanding its articulatory underpinnings and the extent to which they may be considered language-specific versus cross-linguistically applicable.

The articulatory-kinematic underpinnings of PBL, arising with a local slowing down near the boundary, are reflected primarily on the temporal dimension (e.g., Edwards et al., 1991; Byrd & Saltzman, 1998; Cho, 2006), although it can be accompanied by spatial expansion (Byrd & Saltzman 2003; Cho, 2005; Cho, 2006; Li, Kim & Cho, 2023). On the other hand, prominence (or ‘stress’ in a broad term) is likely to induce hyperarticulation of some sort. It typically brings about expansion of articulation in both spatial and temporal dimensions, characterized as being larger in displacement, longer in duration, and faster in movement peak velocity (e.g., de Jong, 1995; Fowler, 1995; Cho, 2006). The prominence-induced strengthening effect is often called *localized hyperarticulation*, as it is generally localized to a stressed syllable. This is in contrast with a communicatively driven hyperarticulation that is assumed to affect the whole utterance in the sense of Hyper- and Hypo-articulation (H & H) theory (Lindblom, 1990). A division between the boundary-induced and prominence-induced effects appears to be particularly relevant for a so-called ‘head-prominence’ language such as English. In a head-prominence language, as discussed in Jun (2014), when a phrase-level stress (e.g., a nuclear pitch accent) falls on a particular word, a lexically stressed syllable of the word becomes the head of the phrase, so that hyperarticulation is localized to the head (i.e., the stressed syllable), lending prominence. In a head-prominence language, therefore, the location of the prominence, being localized to

a stressed syllable, is independently determined from a prosodic boundary, and its function is differentiated from that of marking prosodic boundaries (see Keating, 2006; Shattuck-Hufnagel & Turk, 1996 for related discussion).

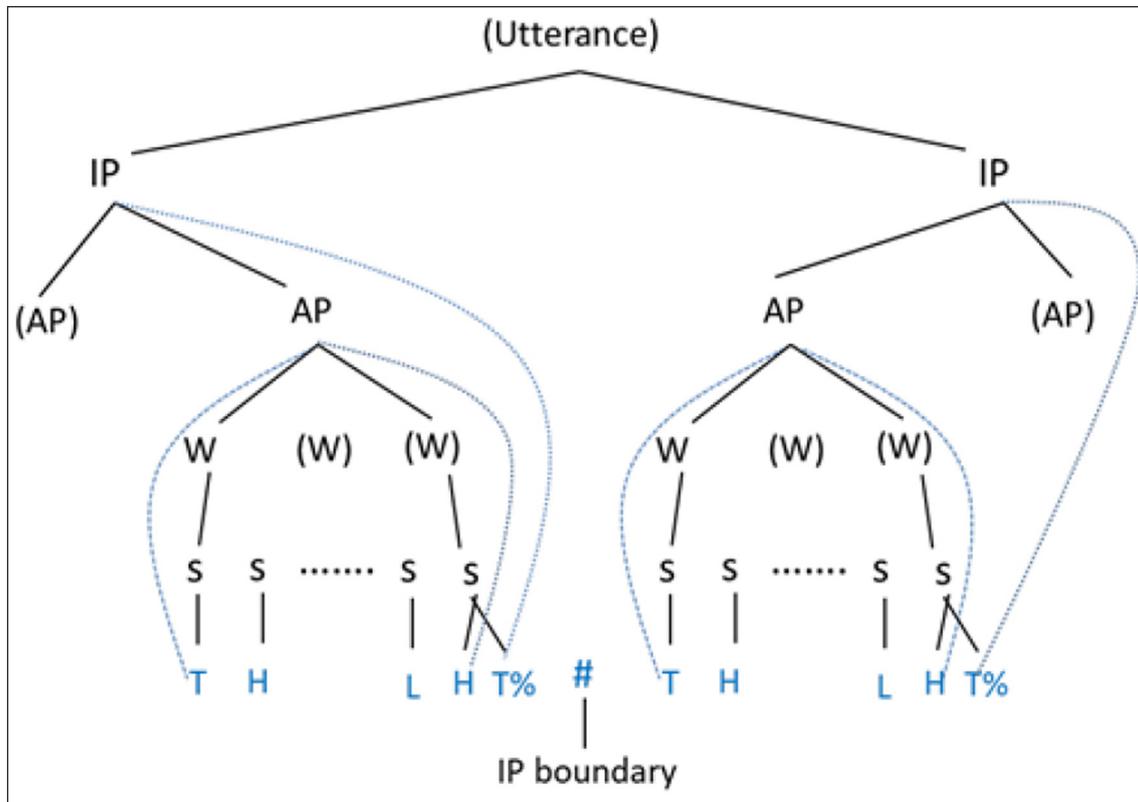
Another prominence-related typology may include languages that carry stress on the final syllable. In such languages that are common in the world languages, as discussed in Gordon (2016), the final syllable may serve as a locus of prominence while simultaneously being the right edge of a phrase when occurring phrase-finally, which may also serve a demarcating function. In particular, Jun (2014) cites French and Kiche as examples of languages that maintain prominence in the phrase-final position through a form of pitch accent, which also functions as the right edge of a phrase, specifically an Accentual Phrase. These languages are referred to as ‘head/edge-prominence’ languages (see Jun, 2014, for other cases that also fall within this category). (Note, however, that whether the localized hyperarticulation due to stress in these languages is indeed further augmented when it aligns with the right edge of a larger phrase remains an empirical question, which falls beyond the scope of the present study.)

In contrast, Korean does not employ lexical stress, fixed stress assignment, or pitch accent in its prominence system. Instead, prominence that may arise with, for example, information or syntactic structure in Korean is typically marked by phrasing or grouping of prosodic constituents, such as a prosodic word, into a larger prosodic unit. In Korean, this prosodic unit is called an Accentual Phrase (Jun, 2014). Such phrasing made by an Accentual Phrase demarcates both *edges* of a phrase with a particular tonal pattern (e.g., a phrase begins and ends with an LH tone). For this reason, Korean is often referred to as an ‘edge-prominence’ language, which is different in prosodic typology from a ‘head-prominence’ language such as English as described above (Jun, 2014). On this account, Jun defines the term ‘edge-prominence’ by considering that prominence is expressed primarily through phrasing, with the beginning and end of a phrase demarcated by edge tones on both sides. Jun’s classification, however, does not specifically elaborate on whether edges, either at the beginning or end of a phrase, may also be accompanied by segmental hyperarticulation in addition to prominence realized through edge-marking tones. Nevertheless, there is evidence suggesting that the left edge (i.e., the beginning of a phrase) may exhibit hyperarticulation to some extent, falling under the concept of ‘domain-initial strengthening.’ In other words, while segments tend to undergo strengthening at phrase-initial positions across languages (e.g., Keating, Cho, Fougeron, & Hsu, 2023; Cho, 2016), Korean exhibits a more pronounced form of boundary-related articulatory strengthening at phrase-initial positions than many other languages (Cho & Keating, 2001; Keating, Cho, Fougeron, & Hsu, 2003). Keating et al. (2003) suggest that this robust domain-initial strengthening effect in Korean aligns with the idea that, in Korean, prominence is primarily conveyed through phrasing. This implies that domain-initial strengthening may serve a dual function, enhancing both prominence

and boundary marking, and establishing the left edge of a phrase as the locus of prominence in both tonal and segmental dimensions. The latter of these dimensions results in a form of hyperarticulation in Korean (e.g., Cho & Jun, 2000).

More generally, in languages where both edges of prosodic constituents (e.g., initial and final positions of a phrase) may become loci of prominence, at least as expressed through intonational means (as opposed to a stressed syllable in head-prominence languages), boundary marking (phrasing), which delineates the edges of prosodic constituents, can be considered commensurate with prominence marking (or lending prominence), possibly accompanied by a form of hyperarticulation. In this view of Korean as an edge-prominence language, we hypothesize that articulatory patterns at the ‘right’ edge of a prosodic constituent, associated with phrase-boundary lengthening (PBL) will also exhibit some degree of hyperarticulation in the segmental dimension such as spatial expansion to a level that is not normally observed at the right edge in a head-prominence language (e.g., English). That is, the predicted right-edge effects are likely to differ from the patterns observed in English, considered a head-prominence language. In English, hyperarticulation is primarily associated with stress-related prominence (e.g., Barnes, 2002), and the phrase-final position typically does not exhibit the same kind of hyperarticulation seen with stress. The present study elaborates on this possibility by investigating preboundary lengthening (PBL) and other kinematic characteristics of articulatory gestures at prosodic boundaries in Korean.

Before addressing the specific research questions in this study, it is important to clarify the relevant levels of phrasing in Korean and specify the level to be tested. **Figure 1** shows a schematic representation of phrasing in an utterance, featuring two Intonational Phrases (IPs), each of which can contain one or more Accentual Phrases. In Jun’s discussion (2014) of phrasing to characterize Korean as an edge-prominence language, she refers to the formation of an Accentual Phrase to express prominence. However, based on the strict layering hypothesis (Selkirk 1984, 1995; also discussed by Shattuck-Hufnagel and Turk, 1996), we assume that Accentual Phrases in Korean are embedded within an Intonational Phrase (IP), aligning an edge of the IP with an edge of an Accentual Phrase (AP). Moreover, as discussed by Keating et al. (2003) and Cho and Keating (2001), a cumulative left-edge effect of domain-initial strengthening typically results in the beginning of an IP boundary showing a more robust edge effect than that of an AP, still demonstrating the characteristics of an edge-prominence language at the left edge. Similarly, in the present study, we investigate the right-edge effect of an IP boundary aligned with that of the AP with the aim of exploring the right-edge effects in both temporal and spatial dimensions in the context (being IP-final) where these effects are expected to be more pronounced.



**Figure 1:** A schematic representation of phrasing in an utterance. The prosodic structure beneath each Intonational Phrase (IP) is adapted from Jun (1993, 2000). Please note that at the beginning of an Accentual Phrase (AP), ‘T’ refers to either ‘H’ or ‘L,’ depending on the onset consonant (with ‘H’ for fortis or aspirated consonants and ‘L’ for others). ‘T%’ refers to a boundary tone associated with the end of an Intonational Phrase (IP).

The primary purpose of the present study is therefore to explore kinematic characteristics of articulation of bisyllabic words (CV.CV and CV.CVC) in relation to PBL in Seoul Korean, so that it adds to the body of cross-linguistic studies on kinematic characteristics of articulation at prosodic junctures. The obtained results will allow us to explore to what extent the kinematic characteristics associated with PBL show a general cross-linguistic tendency versus language-specificity of Korean. We will discuss the results with some theoretical considerations as below.

The first consideration concerns the scope of PBL (i.e., the extent to which PBL in Korean can spread to the left of the final syllable in bisyllabic words). The scope of PBL may be language-specifically determined, but it is also known to be influenced by factors related to the lexical prominence of the language (e.g., Turk & Shattuck-Hufnagel, 2007; Katsika, 2016; Seo et al., 2019). For example, prominence arising with lexical stress may attract PBL toward a non-final

syllable in a head-prominence language like English or Greek, though the exact scope of PBL may differ between the two languages (Katsika, 2016). In an articulatory study, Jang and Katsika (2020) also explored this scope-related question in Seoul Korean (i.e., to what extent PBL may spread to the left into non-final syllables) by examining constriction formation duration and release duration of the consonantal gestures of polysyllabic words at prosodic boundaries in Korean. Their results showed that PBL was largest in magnitude for the final coda consonant, and it was substantially attenuated for the onset consonant of the final syllable, showing a general progressive effect (i.e., progressively decreasing from the right edge). They, however, showed no further evidence on the leftward spreading of PBL beyond the onset of the final syllable, while there was some degree of shortening of consonantal gestures of the penultimate syllable. They also demonstrated that the presence of narrow focus either on the preceding word or on the target word did not influence the leftward spreading (scope) of PBL, showing an independence of PBL from the focus-related prominence.

The present study further elaborates on the PBL in Seoul Korean in the following aspects that complement the previous work by Jang and Katsika (2020). First, while Jang and Katsika examined distribution of PBL on a word whose final syllable is closed (i.e., with a coda), the present study examines whether the scope of PBL can be constrained by the phonetic content of the final syllable (presence or absence of the coda). As discussed in Turk & Shattuck-Hufnagel (2007) and Seo et al. (2019), there is a possibility that PBL operates phonologically on the basis of the syllable (rather than the segment's proximity to the boundary), so that leftward spread will not be constrained by the presence of the coda. But these studies were based on acoustic data, so that they did not capture actual kinematic characteristics of underlying articulatory gestures that constitute a syllable. Thus, it may still be the case that the physical distance from the prosodic juncture has an influence on articulatory kinematic measures, so that the leftward spreading of PBL beyond the final syllable may depend on whether there is a coda or not (see below for related discussion).

Second, the present study includes other kinematic measures such as displacement and movement (peak) velocity, as well as time-to-peak velocity (acceleration duration), and investigates how PBL may be related to variation in these kinematic measures. Examining these measures is particularly important to understand the nature of articulatory strengthening (hyperarticulation) in both spatial and temporal dimensions that occurs at the right edge in Korean as an edge-prominence language. Moreover, understanding the relationship between kinematic measures will allow us to consider the kinematic characteristics of preboundary articulation in dynamical terms (Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014). For example, while PBL may be expected to be associated with a lowered movement velocity (as PBL is often assumed to be caused by a slowing down of articulatory movements), the opposite may be true if PBL turns out to be accompanied by substantial spatial expansion which may cause an increase in movement

velocity due to the natural propensity of high correlation between displacement and movement velocity (Munhall, Ostry, & Parush, 1985; Ostry & Munhall, 1985).

Third, this study aims to compare the previously mentioned preboundary effects under two different prominence conditions. The goal is to examine how boundary-related effects may vary across different levels of prominence. Notably, in Korean, it appears that prominence marking does not strongly affect the timing aspect of the scope and magnitude of PBL (Jang & Katsika, 2020). However, since boundary-related effects often interact with prominence in other kinematic measures (e.g., Katsika, 2016; Cho, 2016; Li et al., 2023), it remains to be seen whether boundary effects on preboundary articulation are further influenced by prominence. It's important to distinguish between different types of prominence in Korean. One type of prominence is hypothesized as edge-related prominence, where both edges of a larger prosodic unit are assumed to carry prominence relative to those of a smaller one, as we also discussed above in the introduction. Another type of prominence may stem from information structural factors, possibly independent of phrasing, which could otherwise contribute to the hypothesized edge-related prominence. In this study, we incorporated 'new information' versus 'given information' contexts related to prominence associated with information structure (cf. Gussenhoven, 2008; Mücke & Grice, 2014). These contexts result in varying levels of prominence, corresponding to 'broad' focus versus 'background.' Notably, we did not include 'narrow' focus in the prominence conditions because it is typically linked to initiating a new phrase in Korean (Jun, 1998; Jeon & Nolan, 2017). This would have added complexity to the interplay between boundary and information structure-related prominence. Hence, the prominence difference observed in the 'new' versus 'given' contexts used in this study may manifest as fine phonetic details distributed throughout the entire utterance, rather than being confined to a specific prosodic unit. This approach allows us to investigate the extent and magnitude of boundary effects in both temporal and spatial dimensions across different levels of prominence that may extend broadly across the entire utterance.

Another important theoretical consideration in the present study pertains to how results to be obtained can be accounted for in dynamical terms, especially by the theory of  $\pi$ -gesture (Byrd & Saltzman, 2003; see Byrd & Krivokapić, 2021 for a related review). The theory of  $\pi$ -gesture (the prosodic gesture) accounts for phonetic variation at prosodic boundaries within the frameworks of the task dynamic model and Articulatory Phonology (e.g., Saltzman & Munhall, 1989; Goldstein, Byrd, & Saltzman, 2006). The  $\pi$ -gesture is different from a usual 'tract variable' (constriction) gesture, which is realized with a vocal-tract constriction (in terms of its degree and location). It is a "non-tract variable" gesture (with no specification of constriction degree and location), assumed to be anchored at a prosodic boundary. Crucially, it modulates the rate of the "clock" that controls articulatory temporal activation of constriction gestures in the vicinity of the prosodic juncture. Thus, the temporal expansion that occurs at prosodic junctures (at both

edges of a prosodic constituent) does not stem directly from settings of the dynamical parameters such as the target and the stiffness (cf. Saltzman & Munhall, 1989), but it comes about as a consequence of a slowing-down of the clock that is modulated by a  $\pi$ -gesture. The  $\pi$ -gesture is governed by the prosodic constituency so that the degree of influence of the  $\pi$ -gesture is determined by boundary strength (the size of the boundary): The larger the prosodic constituent, the stronger its effect, resulting in larger temporal expansion. Its influence is strongest at the juncture and becomes gradually attenuated as it gets farther away from the juncture. This accounts for the progressively decreasing magnitude of PBL from the right edge to the left, found across languages. However, if, as hypothesized, the right edge in Korean, as an edge-prominence language, indeed undergoes hyperarticulation, showing a robust spatial expansion comparable to the hyperarticulation effects possibly beyond what could be expected from the influence of the  $\pi$ -gesture, it would be valuable to discuss how the potential strengthening of preboundary articulation aligns with the operation of the  $\pi$ -gesture at prosodic junctures in Korean as an edge-prominence language.

## 2. Methods

### 2.1. Speech materials

Eight bisyllabic target words were used for production: /mami/, /mima/, /p\*ap\*i/, /p\*ip\*a/, /mamim/, /mimam/, /p\*ap\*ip/, and /p\*ip\*ap/ (here /p\*/ refers to the Korean fortis (tense) stop; cf. Cho, Jun & Ladefoged, 2001). These words were introduced as pet names in the mini dialogue shown in **Table 1**. These words included two bilabial consonants (/m/ or /p\*/) to examine the variation in lip movement, and two vowel sequences (/CaCi/ or /CiCa/) to factor in the influences of vertical tongue movements (in /a/-to-/i/ or /i/-to/a/). The final syllable was open or closed (CV.CV or CV.CVC) to examine the influence of the syllable structure on scope of PBL (i.e., whether PBL spreads to the left on a syllable basis or on a segment basis).

BND	Info	Example sentences	English translations
(a) Open Syllable Condition: Test word = CV.CV (e.g., /mima/)			
IP-final	'new' (br. foc.)	A: [musin il is*at*ε]? B: [jʌŋmaninε <u>mima</u> ]#[pinu maŋat*ε] soap ate	What happened? Youngman's Mima ate the soap.
	'given'	A: [jʌŋmaninε mima]#[pinu atf*εt*ε]? B: [jʌŋmaninε <u>mima</u> ]#[pinu maŋat*ε] soap ate	What did Youngman's Mima do with the soap? Youngman's Mima ate the soap.

(Contd.)



Boundary (IP-final/IP-medial) and Info-Structure ('new'/'given') were pivotal experimental factors. As exemplified in **Table 1**, mini discourse contexts were constructed, in which the participant played the role of Speaker 'B' in response to a prompt question ('A'). Each test word occurred in IP-final or IP-medial conditions (Boundary). Orthographic and syntactic schemes were used to guide intended prosodic patterns. IP-final renditions were guided by a comma after the test word, which was aligned with a major syntactic juncture between an NP and a VP<sup>1</sup>. It's important to note that when constructing this structure, we conducted informal preliminary testing with Korean speakers. The results confirmed that in the IP condition, sentences were naturally split into two IPs, with an IP boundary placed in the intended location even without a comma, regardless of the information structural conditions. Nevertheless, to ensure a consistent IP boundary condition across all participants, we added a comma after the intended target word.

IP-medial conditions were established by grouping the test word and the following word with no space in between, forming a noun phrase likely as a noun-noun compound (e.g., /mima pinu/ or /mimam sat<sup>h</sup>aq/). In our informal preliminary testing, speakers generally did not introduce a major phrase boundary between the two nouns once they understood the meaning of the noun phrase. Nevertheless, to facilitate the grouping and ensure consistent phrase-medial conditions across all speakers, we employed the strategy of putting no space between the two nouns. The prosodic boundary between the two nouns that may form a compound may not be considered a prosodic word boundary, although the size of such a boundary in this non-lexicalized compound may be possibly larger than the one in a lexicalized one. In other words, this IP-medial condition aligns with a lexical word boundary, though it may not be parsed as a prosodic word boundary.

Two conditions 'new' and 'given' were included for Info-Structure, which induced two possible levels of relative prominence. The 'new' condition was elicited by the question 'What happened?', whereas the 'given' condition was ensured by the test word having been 'given' in the prompt question, with the contrastive focus (in bold) elsewhere in the test sentence. It is also worth noting here that in the 'new' condition, the focus-induced prominence prompted a broad focus in the answer, which is likely to be weaker than the one arising with contrastive (narrow) focus (Mücke & Grice, 2014; cf. Gussenhoven, 2008).<sup>2</sup>

At this point, it's important to address a concern raised by an anonymous reviewer regarding the validity of the boundary and information structure-related conditions embedded in our experimental stimuli. Specifically, it was pointed out that in the 'new' (broad focus) condition, splitting the sentence into two IPs in response to 'what happened' doesn't align with natural phrasing for broad

<sup>1</sup> We follow the strict layering hypothesis (Selkirk, 1984, 1995; see also Shattuck-Hufnagel & Turk, 1996), so that one or more smaller phrases, which are called Accentual Phrases in Korean, are embedded in an IP. This means that an edge of IP is being aligned with an edge of an AP.

<sup>2</sup> Note that, as can be seen in **Table 1**, in the 'given' condition, due to the nature of the elicitation procedure, a narrow focus occurred somewhere after the target word in response to 'who' or 'what' in the question sentence. It was ensured that the context word immediately following the target word did not receive a narrow focus in the 'given' condition.

focus. Consequently, the IP boundary, as the reviewer argues, may not be considered as natural as one might expect in the broad focus condition. This criticism is rooted in the assumption that such splitting redistributes prominence-lending units across two IPs, whereas a single IP would be expected in the broad focus condition. While theoretically, such split IPs may function as independent prominence-lending units, especially when each unit includes the head of the prominence as is often assumed in a head-prominence language, this does not necessarily mean that speakers do not produce an IP boundary in a broad focus utterance. There are several reasons for this.

First, both our pilot testing and our native intuitions as trained Korean prosodic experts suggested that there would likely be an IP boundary in the broad focus condition, resulting in neutral renditions. Second, speakers frequently insert an IP boundary between an NP and a VP in a neutral context (as discussed in Shattuck-Hufnagel & Turk, 1996). Third, we are not aware of any experimental studies that provide solid evidence indicating that broad focus must always be realized with a single IP. Fourth, despite the possibility that such an IP could stem from other confounding factors that may influence phrasing, at the very least, in the ‘new’ (broad focus) context, the target word did not receive either a narrow focus or a contrastive focus. In this sense, the level of prominence in the IP-final condition of the ‘new’ context aligns with the ‘broad’ focus category. Lastly, and, more importantly, in accordance with Beckman (1996) and Shattuck-Hufnagel and Turk (1996), we assume that prosodic structure is a grammatical entity and prosodic structuring is autonomous, being parsed on its own. Consequently, we posit that the differences observed between IP-final and IP-medial boundaries are directly conditioned by prosodic structure, even though prosodic phrasing may ultimately be determined by the combined effects of various factors influencing prosodic phrasing.

For these reasons, many experimental phonetic studies on preboundary effects have deliberately employed different syntactic structures or other structural means to induce various prosodic structures. This approach allows for the testing of prosodic boundary effects, even if different phrasings result from potentially confounding factors (e.g., Byrd & Saltzman, 1998; 2003; Byrd et al., 2000; Jang & Katsika, 2020; Katsika, 2016; Keating, Cho, Fougeron, & Hsu, 2003; Cho & Keating, 2001, 2009; Krivokapić, Styler, & Parrell, 2020). However, this does not imply that an IP must be realized with the same phonetic content. As pointed out by the reviewer, it’s possible that an IP may exhibit varying strength depending on the contributing factors, resulting in different phonetic effects (cf. Ladd, 2008; Krivokapić & Byrd, 2012). Nevertheless, testing this possibility falls outside the scope of the present study, which primarily aims to initiate an investigation into the effect of prosodic boundaries on the kinematic realizations of preboundary words in Korean, an area that has not been previously explored.

A total of 1,152 tokens were collected (8 words × 2 boundary conditions × 2 Info-Structure conditions × 4 repetitions × 9 speakers). Due to measurement-related errors (e.g., uncertainty of pinpointing kinematic landmarks), five tokens were discarded in all analyses and 24 additional tokens were discarded in the time-to-peak velocity analysis.

## 2.2. Apparatus and procedure

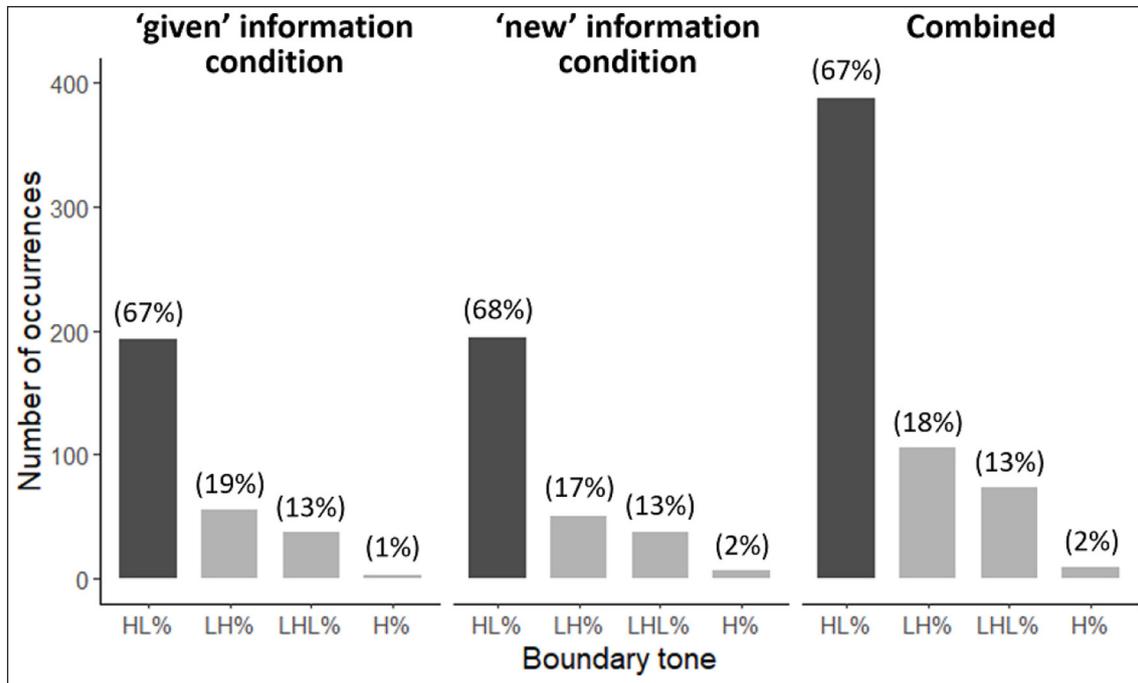
Articulatory data were collected from nine native speakers of Seoul Korean (five male and four female college students in their 20's) using 2D Electromagnetic Midsagittal Articulography (Carstens AG200). Sensors were attached to the upper and lower lips (and on the front/back of the tongue). There was a practice session prior to the experiment where participants were familiarized with all the test words appearing in the mini discourse contexts as names of pet dogs along with pictures. During the experiment, participants heard a pre-recorded prompt question and read the test sentence as a corresponding answer shown on a computer screen. Trials were blocked by test words with different randomized orders over four repetitions. Participants made errors in some trials, the majority of which consisted of the choice of unintended prosodic rendition or disfluency due to the sensors and wires. Whenever an error was spotted by the experimenter (trained on the Korean prosodic transcription), the participant was asked to read the sentence again. Prosodic renditions (including boundary types, IP-final vs. IP-medial) were cross-checked by two other trained Korean phoneticians (the second and third authors), who confirmed that all recorded tokens were naturally produced with the intended prosody.

Furthermore, all three authors transcribed the tones on the target words to examine the types of boundary tones in our experimental sentences. In the IP-medial condition, the tone consistently realized on the final syllable of the target word was 'H', which occurs in the middle of an Accentual Phrase embedded inside an IP. On the other hand, in the IP-final conditions, various boundary tones were observed. **Figure 2** illustrates IP-final boundary tones in both the 'given' and 'new' conditions, as well as their combined representation. Notably, a complex HL% (falling) tone is the most frequent (67%) in the IP-final condition, followed by another complex LH% (rising) tone (18%) and a rise-fall LHL% (13%). Only 2% of the IP conditions were produced with an H tone. The distribution of boundary tones is notably similar across both the 'given' and 'new' conditions.<sup>3</sup> Our data, as shown in **Figure 2**, consistently shows that a complex tone, which may be realized over a longer temporal extent than a simple tone (cf. Zhang, 2014), is prevalent in the majority of IP-final tokens (98%).

Speaker variation in the choice of boundary tones can also be summarized as below. Notably, while there were some individual variations in choosing different boundary tones, some

---

<sup>3</sup> Note that we attempted to perform statistical modelling with tonal type as an additional predictor, especially considering that, among the complex tones, rising tones are generally expected to have a longer temporal extent than falling tones (e.g., Ohala & Ewan, 1973; Myers, 2003; Kentner, Franz, Knoop, & Menninghaus, 2023; Li, Kim, & Cho, 2023). However, we decided not to report the results for the following reasons. First, the returned results were too complex to understand, presumably due to the unbalanced distribution of tokens across different tonal types. Second, the effects of the tonal types and their interactions with various other experimental conditions were beyond the scope of our original research questions. Therefore, we have deferred the investigation of tonal type effects in conjunction with preboundary phonetic effects to future studies.



**Figure 2:** Types of IP-final boundary tones in the 'given' and 'new' conditions, and pooled across the two conditions.

consistency was also observed. In particular, seven speakers (out of nine) produced a falling tone (HL%) quite consistently, which was the most frequent boundary tone.

- HL% (67%): Seven out of nine speakers (S01-S02, S05-S09) utilized the most frequent HL% (falling) tone, accounting for 95% of all HL% (falling) tokens (369 out of 388).
- LH% (18%): The remaining two speakers (S03, S04) consistently produced LH% (rising tone), accounting for 97% of all LH% tokens (102 out of 105).
- LHL% (13%): Two speakers (S01, S09) generated 75% of the LHL% (54 out of 74), while the remaining 25% (20 tokens) were produced by five other speakers (S02-S06).

### 2.3. Measurement and statistical analyses

The movement data for lip closing and opening were obtained from the Euclidean distance of the two sensors on the upper and lower lips (i.e., Lip Aperture). Duration (DUR), time-to-peak velocity (T-to-PKVEL), peak velocity (PKVEL), and displacement (DISP) were obtained using Mview (Tiede, 2005; cf., Cho, Son, & Kim, 2016). See **Figure 3** in the results section for schematized measures. The onset and target of the gesture were defined as time points at 20% of PKVEL (mm/s) during acceleration (for the onset) and deceleration (for the target). DUR (in ms) was measured from the onset of the gesture under investigation to the onset of

the following gesture (including the plateau).<sup>4</sup> Time-to-PKVEL (in ms) was measured from the onset to the attainment of peak velocity, which is roughly the same as acceleration duration. This durational measure is considered to reflect the temporal control of the clock-slowness rate by the pi-gesture (or the gestural stiffness as a dynamical parameter) more accurately than the entire movement duration because the second component of the gesture after the peak velocity attainment (roughly the same as the deceleration duration) is subject to truncation due to an earlier activation of the following gesture (Byrd & Saltzman, 2003). DISP (in mm) was the spatial change (in Lip Aperture) from the onset to the target.

With regards to the examination of spatio-temporal change in Lip Aperture for CV.CV(C) words, a theoretically-related caveat is in order. Given that all the consonants are bilabial, the lip closing movement is directly relevant to the consonantal gesture. But as for the lip opening movement, one could assume that it is related to the vowel since it is aligned with the opening of the vocal tract (which is also proximally aligned with the onset of the vowel in the acoustic dimension). But in the framework of Articulatory Phonology, some researchers suggested that a consonant could be modeled as having two gestural components (i.e., the closing gesture and the release gesture, as proposed in the split-gesture dynamics model by Nam, 2007), suggesting that the opening movement may be associated with the consonantal gesture. (But see Iskarous & Pouplier, 2022 for comments on possible theoretical issues related to the split-gesture dynamics model.) However, even in a split-gesture model, once the consonantal closure is released as specified by the consonantal release gesture, the lip opening movement continues beyond an assumed equilibrium position, in correlation with the vocalic movement. This continuation may be in part due to the influence of jaw movement, which accompanies tongue movement and affects lip opening. Consequently, the continued lip opening movement after the release gesture can no longer be attributed solely to the activation of the consonantal release gesture. It is therefore plausible that the later part of the lip opening movement into the vowel is constrained by the vocalic gesture and could possibly be seen as a proxy for the vocalic gesture. Taking all of this into consideration, we suggest that lip closing and opening movements are related to CV articulation, though the precise modeling of the relationship between the lip opening gesture and the vocalic gesture in dynamical terms remains to be explored. It is also worth noting that Jang and Katsika (2020) examined the consonantal closing and opening movements to assess PBL effects.

It should also be noted here that for the final coda in CV.CVC, one might think that the rightmost articulatory component that is immediately adjacent to a prosodic boundary must be

---

<sup>4</sup> Note that this kinematically-defined durational measure excluded a pause in case one occurred in the IP-final condition, allowing us to consider preboundary lengthening as a pre-pausal effect. While our analysis did not account for the potential influence of a pause on pre-pausal lengthening, the preboundary phenomenon was typically accompanied by a complex boundary tone, signifying its role as phrase-final (pre-pausal), often coinciding with  $f_0$  information that collectively signals prosodic juncture.

the coda's opening movement associated with its constriction release. This was pointed out by an anonymous reviewer, rightly suggesting that PBL for the closed syllable word of CV.CVC could not be adequately captured without examining the temporal realization of the coda's opening component. But this is not the case in Korean. In Korean, there exists a phonological rule whereby a coda obstruent is never released unless it is resyllabified as an onset when followed by a vowel-initial syllable within the same phrase. Due to this phonological rule, speakers often maintain the closure for such an extended period in an utterance-final position that it cannot be considered as being a phonologically-specified release component of the gesture. The phonological constraint that does not license the release in the coda can be translated into gestural terms, especially within the framework of Articulatory Phonology (Goldstein et al., 2006). In such a case, the timing of the release component of a gesture is not specified in the gestural activation of the coda consonant. In the IP-medial context of the current study, the target word in the closed syllable context (e.g., /p\*ap\*ip/) was indeed followed by another word that began with a fricative /s/ (e.g., /sat<sup>h</sup>aŋ/), preventing it from being resyllabified as an onset. Consequently, the release of the coda is not part of the gestural component of the target word. Instead, it is triggered by the following /s/.

When separated by an IP-final boundary, the constriction may eventually be released, but the timing and displacement of the release, as it is not specified, vary significantly. This release may be initiated by a tendency to return to a rest position before starting a new IP or in preparation for the articulation of the following word. In fact, our kinematic data indicate that the release of the coda in the IP condition is extremely variable, such that it is often delayed until the beginning of the new IP, even in the presence of some pause in both oral and nasal stop conditions. Thus, in practice, we could not reliably measure the release component of the coda consonant, and in theory, we assumed that the release component under consideration was not part of the gestural component being specified. For these reasons, we considered the closing component of the gesture (rather than the release component) as the last measurable articulatory component of the final C in CV.CVC.

Statistical analyses were carried out with R4.0.5, fitting a linear mixed-effects (LME) model to raw values of each measure (DUR, T-to-PKVEL, DISP, and PKVEL) for each closing and opening gesture.<sup>5</sup> The model included four binary variables as fixed effects, all of which were deviation-coded with the underlined level as the reference level. Boundary (IP-medial or IP-final) was

---

<sup>5</sup> Given that the length of the Intonational Phrase (IP) containing the target word varied across the Boundary levels (IP-medial = 8 syllables, IP-final = 6 syllables), we received advice to normalize the two temporal measures (DUR and Time-to-PKVEL) according to the IP length. However, our primary analysis utilized raw values, assuming that preboundary lengthening (PBL) is localized primarily in IP-final syllables, independently of the global effect of IP length. It appears that this same assumption guided previous studies on PBL (e.g., Byrd & Saltzman, 1998; Byrd & Saltzman, 2003; Katsika, 2016), which also examined speech materials of varying IP lengths across boundary conditions without normalizing the measures.

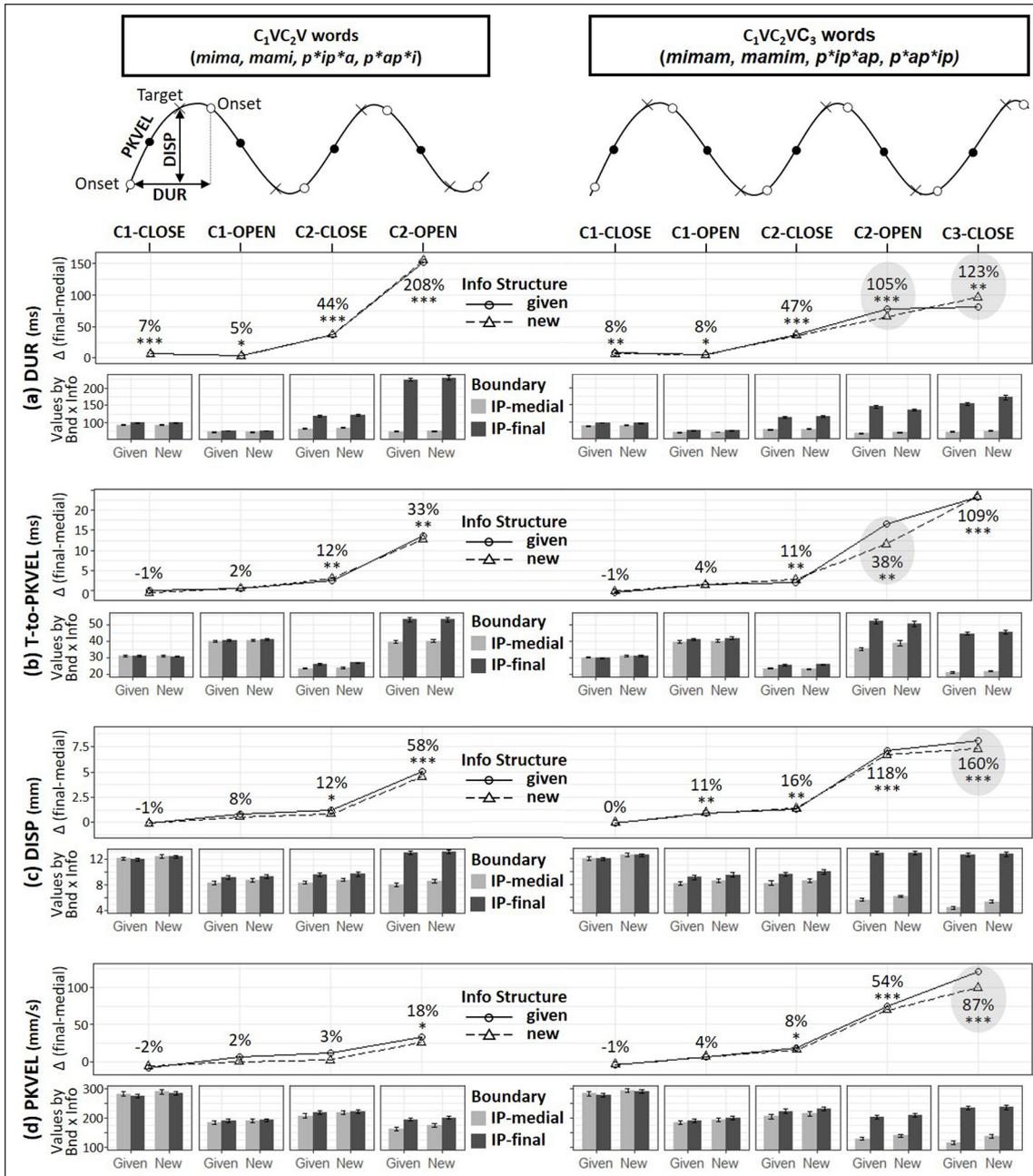
included as the main experimental predictor. Info-Structure (given or new), Consonant (/m/ or /p\*/), and Vowel-sequence (/a-i/ or /i-a/) were also included as control predictors. For the purpose of the present study, only two-way interactions between Boundary and each control predictor were included as fixed effects to avoid overfitting. Initially, we attempted to include the maximal random effects structure for each model we built, but models did not converge most of the time. Thus, we conducted likelihood ratio tests to trim down effects that did not reach significance or induced non-convergence, beginning from the maximal random effects structure justified by the design (i.e., all by-participant and by-item intercepts and slopes for Boundary, Info-structure, and their interaction). As a result, by-participant and by-item intercepts and by-participant slope for Boundary remained in the final model structure. (See Appendix 1 for detailed results of LMEMs and R syntaxes used.)

Finally, we will also probe into relationships in variation between kinematic measures of duration, displacement, and peak velocity. These additional analyses will allow us to understand how much variation in temporal dimension (to be reflected in PBL) can be accounted for by variation in spatial dimension (to be reflected in variation in displacement), and vice versa.

### 3. Results

#### 3.1. Boundary effects on kinematic measures

Boundary effects on each of the four measure types (DUR, T-to-PKVEL, DISP, and PKVEL) are plotted in two sets of graphs in **Figure 3**. First, the bar graphs show the mean values obtained from each boundary level of each closing and opening gesture, separately for the Info-Structure levels (new or given). Second, line graphs are provided for simplified presentation of the magnitude of the boundary effects, visualizing  $\Delta$ (final-medial) values (i.e., the mean difference between the boundary levels) in each level of Info-Structure, so that a larger  $\Delta$  value indicates a larger boundary effect. The results of LME models for the main effects of Boundary and Info-Structure are summarized in **Table 2**. See Appendix 1 for the results of other control predictors included in the model (i.e., Consonant sequence, Vowel sequence, their interactions with Boundary, and interaction between Boundary and Info-Structure). In **Figure 3**, significant Boundary effects are indicated by 3 levels (\* < .05, \*\* < .01, \*\*\* < .001) for the sake of information, but with no *a priori* implication that *p*-values below 0.05 can be used for assessing the robustness or the size of the significance. The figure (and **Table 2a–b**) also contains %-increase, indicating the mean relative increase in each measure from IP-medial to IP-final condition. In the presentation of the results, gestures labelled ‘C1-closing’ and ‘C1-opening’ are associated with C1 in  $C_1VC_2V(C_3)$ , and those labelled ‘C2-closing’ and ‘C2-opening’ with C2 in  $C_1VC_2V(C_3)$ . For C3 in the closed syllable context, only ‘C3-closing’ is included as explained in the method section.



**Figure 3:** Effects of Boundary. The bar graphs show raw mean values with the error bars representing standard errors of the mean. Δ values and %-increase values in the line graphs show the mean increase from IP-medial to IP-final. The significance notations (\* < .05, \*\* < .01, \*\*\* < .001) refer to the Boundary effects. Ovals in gray refer to cases when there is a significant interaction between Boundary and Info-Structure ('new' vs. 'given').

Word type	CV.CV				CV.CVC				
Gesture ID	C1-closing	C1-opening	C2-closing	C2-opening	C1-closing	C1-opening	C2-closing	C2-opening	C3-closing
(a) Boundary (IP-medial = ref.)									
DV = DUR	6.59 (7%) <.001	3.57 (5%) .046	36.90 (44%) <.001	153.32 (208%) <.001	7.35 (8%) .002	5.32 (8%) .013	36.41 (47%) <.001	71.46 (105%) <.001	89.17 (123%) .001
DV = T_PV	-0.20 (-1%) .745	0.72 (2%) .384	2.85 (12%) .002	13.07 (33%) .001	-0.25 (-1%) .653	1.57 (4%) .101	2.45 (11%) .002	14.14 (38%) .001	23.31 (109%) <.001
DV = DISP	-7.09 (-1%) .801	69.43 (8%) .064	103.27 (12%) .024	482.64 (58%) <.001	-1.81 (0%) .932	94.74 (11%) .009	136.44 (16%) .001	698.00 (118%) <.001	774.38 (160%) <.001
DV = PKVEL	-6.05 (-2%) .433	3.13 (2%) .603	7.19 (3%) .364	29.75 (18%) .016	-3.42 (-1%) .577	6.65 (4%) .288	17.19 (8%) .043	72.65 (54%) <.001	110.76 (87%) <.001
(b) Info-Structure ('given' = ref.)									
DV = DUR	0.11 (0%) .879	0.06 (0%) .931	3.32 (3%) <.001	3.35 (2%) .223	0.75 (1%) .337	0.29 (0%) .639	1.85 (2%) .128	-3.89 (-4%) .021	11.09 (10%) <.001
DV = T_PV	-0.05 (0%) .903	0.67 (2%) .157	0.61 (2%) .041	0.33 (1%) .699	0.89 (3%) .017	0.55 (1%) .313	-0.06 (0%) .813	1.09 (3%) .380	0.65 (2%) .335
DV = DISP	37.34 (3%) .002	27.65 (3%) .034	27.03 (3%) .046	34.52 (3%) .027	55.28 (5%) <.001	38.05 (4%) .006	37.22 (4%) .011	26.03 (3%) .061	46.47 (6%) <.001
DV = PKVEL	8.77 (3%) .001	5.38 (3%) .050	7.53 (4%) .018	9.50 (5%) .001	11.44 (4%) <.001	9.83 (5%) .001	8.52 (4%) .006	8.01 (5%) .002	11.38 (7%) <.001

**Table 2:** A summary of the LME models for effects of Boundary (IP-medial = ref.) and Info-Structure ('given' = ref.). Positive coefficients indicate an increase from the IP-medial to the IP-final condition (Boundary effect) or from the 'given' to the 'new' (broad focus) condition (Info-Structure effect). %-increase is given in parenthesis. Significant effects ( $p < 0.05$ ) are marked in gray.

Duration (DUR, **Figure 3a**) shows  $\Delta(\text{final-medial})$  significantly above '0' in all gestures of CV.CV and CV.CVC words. The distribution of PBL over the bisyllabic words can be characterized as being 'progressive' in that the magnitude of PBL (estimated by the %-increase and  $\beta$  in **Table 2**) is progressively decreasing from the right edge. Notably,  $\Delta(\text{final-medial})$  in DUR decreases with a substantial drop from C2-opening to C2-closing as the gesture becomes distal from the boundary. As can be seen by comparing the left and the right panel in **Figure 3a**, such a drop was much steeper in the open syllable condition where C2-opening was the right most gesture, than in the closed syllable condition where C2-opening was not the rightmost gesture. It was also noticeable that even the most distal C1-closing gesture and the following C1-opening gesture in both  $C_1VC_2V$  and  $C_1VC_2VC_3$  show a significant PBL effect on the movement duration (DUR), though with a more attenuated magnitude. This suggests that PBL in Korean can extend to the leftmost closing gesture in bisyllabic words.

Another durational measure Time-to-PKVEL (acceleration duration), however, indicates more restricted distribution of PBL. As can be seen in **Figure 3b**, it shows a significant boundary effect only on those gestures that belong to the final syllable whether open or closed, and there is no distal effect on the gestures (C1-closing, C1-opening) of the first (non-final) syllable. (This will be discussed in relation to the scope of the temporal modulation by the  $\pi$ -gesture in the discussion section.) Within the final syllable, a progressive decrease of the magnitude in duration from the right edge is also evident in Time-to-PKVEL, with a substantial drop from C2-opening (into the following vowel) to C2-closing (from the preceding vowel), though in a lesser degree.

Displacement (DISP, **Figure 3c**) and Peak velocity (PKVEL, **Figure 3d**) also show a progressively decreasing trend from the right edge, though not clearly to the leftmost distal gestures. As can be seen in **Figure 3c**, the Boundary effect on DISP is not significant in initial gestures (i.e., there was no effect on C1-closing and C1-opening in  $C_1V.C_2V$ , and C1-closing in  $C_1V.C_2VC_3$ ). But  $\Delta(\text{final-medial})$  in DISP decreases (again with a substantial drop from C2-opening to C2-closing) as the gesture becomes distal from the boundary. These displacement results indicate that PBL is accompanied by substantial spatial extension. As for PKVEL, the significance of  $\Delta(\text{final-medial})$  is limited to gestures that are proximal to the boundary (i.e., only at the final gesture (C2-opening) in  $C_1V.C_2V$  and at the three gestures that belong to the final syllable (C2-closing, C2-opening, C3-closing) in  $C_1V.C_2VC_3$ ) with relatively smaller %-increase values compared to the effect on DISP. Thus, the PKVEL results show faster articulatory movements at the IP-final position compared to the IP-medial position for proximal gestures near the prosodic juncture, although there is not a clear correspondence between the variation in PKVEL and the durational measure (DUR) of PBL.

When examining the bar plots, one can notice a difference in the nature of the progressively decreasing trend from the right edge between the spatial measure (DISP) and the temporal measure (DUR). As can be seen in the lower panels (bar plots) of **Figure 3a**, DUR values in IP-medial conditions consistently stay low throughout the entire target word. In contrast, in IP-final conditions, DUR values gradually increase toward the right edge, illustrating the waxing boundary effect. DISP exhibits somewhat different patterns. Notably, in the first gesture (C1-closing gesture), the overall DISP values are relatively larger in both IP-medial and IP-final conditions compared to the following gestures within the target word, possibly indicating a word-initial strengthening effect. Interestingly, in the IP-medial condition, the degree of DISP continues to decrease toward the end of the target word, indicating a continuous phrase-medial reduction. Conversely, in the IP-final condition, the degree of DISP begins to increase from the second gesture (C1-opening gesture) onwards, reaching its maximum displacement at the phrase-final gesture, showing articulatory strengthening on the right edge. However, when considering the mean differences ( $\Delta$ ) presented in the upper panels of **Figure 3c** (and the  $\beta$  values provided in **Table 2**), the progressively increasing trend in DISP towards the right edge still appears to originate from the initial C1-closing gesture and extend to the right edge. Thus, it appears that this increasing pattern, contributing to significant boundary effects in the spatial dimension, results from both the gradual increase in articulatory strengthening represented by DISP in the IP-final condition and the gradual reduction observed in the IP-medial condition.

These results, taken together, indicate some similarities and differences between CV.CV and CV.CVC. As shown in **Figure 3a**, PBL spreads leftward to a similar extent irrespective of the presence or absence of an additional gesture of C3-closing in  $C_1V.C_2V$  vs.  $C_1V.C_2VC_3$ . In particular, the magnitude of PBL on the C2-closing gesture is comparable (i.e., 44%-increase and 47%-increase in  $C_1V.C_2V$  and  $C_1V.C_2VC_3$ , respectively). On the other hand, the magnitude of the PBL (208%-increase) for C2-opening when it is rightmost in  $C_1V.C_2V$  (i.e., in the absence of C3-closing) is clearly greater than that of either C2-opening (105%-increase) or C3-closing (123%-increase) in  $C_1V.C_2VC_3$ .<sup>6</sup> It appears that the effect of PBL on the final syllable is comparable regardless of whether the final syllable is open or closed. Interestingly, however, Time-to-PKVEL, as can be seen in **Figure 3b**, shows a somewhat different pattern—i.e., the magnitude of PBL reflected in Time-to-PKVEL is comparable on the C2-opening gesture (33%-increase vs. 38%-increase in  $C_1V.C_2V$  and  $C_1V.C_2VC_3$ , respectively), while it was much larger on the rightmost, C3-closing gesture (109%-increase) in  $C_1V.C_2VC_3$  than on the rightmost, C2-opening gesture (33%-increase) in  $C_1V.C_2V$ . Thus, there appears to be some differences in temporal distribution as a function of the syllable structure or whether the rightmost gesture is consonantal opening (C2-opening) versus

---

<sup>6</sup> This observation was further supported by a significant interaction between Boundary and Coda ( $\beta = 81.78$ ,  $p < 0.001$ ) returned by a separate LME model that was fit to the data for C2-opening with Boundary and Coda (presence/absence of C3-closing gesture) as fixed effects.

closing (C3-closing). Nevertheless, what remains invariant is a general pattern of progressive effect (i.e., the more proximal the gesture is to the boundary, the larger temporal expansion the gesture shows in a gradient fashion regardless of whether the final syllable is open or closed).

Another noticeable difference as a function of syllable structure (or whether the final gesture is an opening or a closing one) can be observed with other non-temporal measures (i.e., DISP and PKVEL), which indicate much more pronounced boundary effects on the lip closing movement into the coda consonant in the closed syllable. As can be seen in **Figure 3c–d**, the magnitude of the boundary effect on DISP and PKVEL for the C2-opening gesture is far greater in the closed syllable ( $C_1V.C_2VC_3$ ) context (118%-increase in DISP and 54%-increase in PKVEL) compared to that in the open syllable ( $C_1V.C_2V$ ) context (58%-increase in DISP and 18%-increase in PKVEL). Moreover, the magnitude for each measure for the rightmost gesture (C3-closing) in  $C_1V.C_2VC_3$  (closed syllable condition) is also far greater than for the rightmost gesture (C2-opening) in  $C_1V.C_2V$  (open syllable condition). The rightmost C3-closing gesture in  $C_1V.C_2VC_3$  shows much more boundary-induced spatial expansion (160%-increase in DISP) along with much faster movement (87%-increase in PKVEL) compared to the rightmost C2-opening gesture in  $C_1V.C_2V$  (58%-increase in DISP and 18%-increase in PKVEL). These results indicate that the boundary effect on the gestures that constitute a final syllable is larger in magnitude when the syllable is closed (or when the final gesture involves a closing movement) than it is open.

Additionally, as summarized in **Table 2b** and also seen in **Figure 3a–b**, the effects of Info-Structure on the temporal expansion appear to be much weaker and less systematic as compared to the Boundary effects. The duration measure (DUR) showed a small but significant increase only in two cases (i.e., on C2-closing in the open syllable ( $C_1V.C_2V$ ) condition and on C3-closing in the closed syllable ( $C_1V.C_2VC_3$ ) condition). Similarly, Time-to-PKVEL showed a significant increase only in two cases that are different from the effects on DUR (i.e., on C2-closing in the open syllable ( $C_1V.C_2V$ ) condition and on C1-closing in the closed syllable ( $C_1V.C_2VC_3$ ) condition). Moreover, as can be inferred from the  $\beta$  coefficients and %-increase provided in **Table 2**, even for these sparsely observed significant effects, the magnitude in the focus-induced increase (e.g., about 3% increase in DUR for C2-closing in  $C_1V.C_2V$ ) was far smaller compared to the boundary-related increase in duration (e.g., about 44% increase in DUR for C2-closing in  $C_1V.C_2V$ ).

On the other hand, DISP and PKVEL reveal robust Info-Structure effects across all gestures except for one case (i.e., DISP of C2-opening in  $C_1V.C_2VC_3$  (which showed a marginal effect) as can be seen in **Table 2b**). As can be inferred from %-increase values in the table, the magnitude of change in spatial displacement and movement peak velocity due to Info-Structure appears to be largely similar across all gestures. These results taken together indicate that the prominence associated with broad focus (in the ‘new’ condition) is characterized by spatial expansion accompanied by an increase in peak velocity, but not in duration, and is distributed almost entirely over the bisyllabic words, whether the final syllable is open or closed.

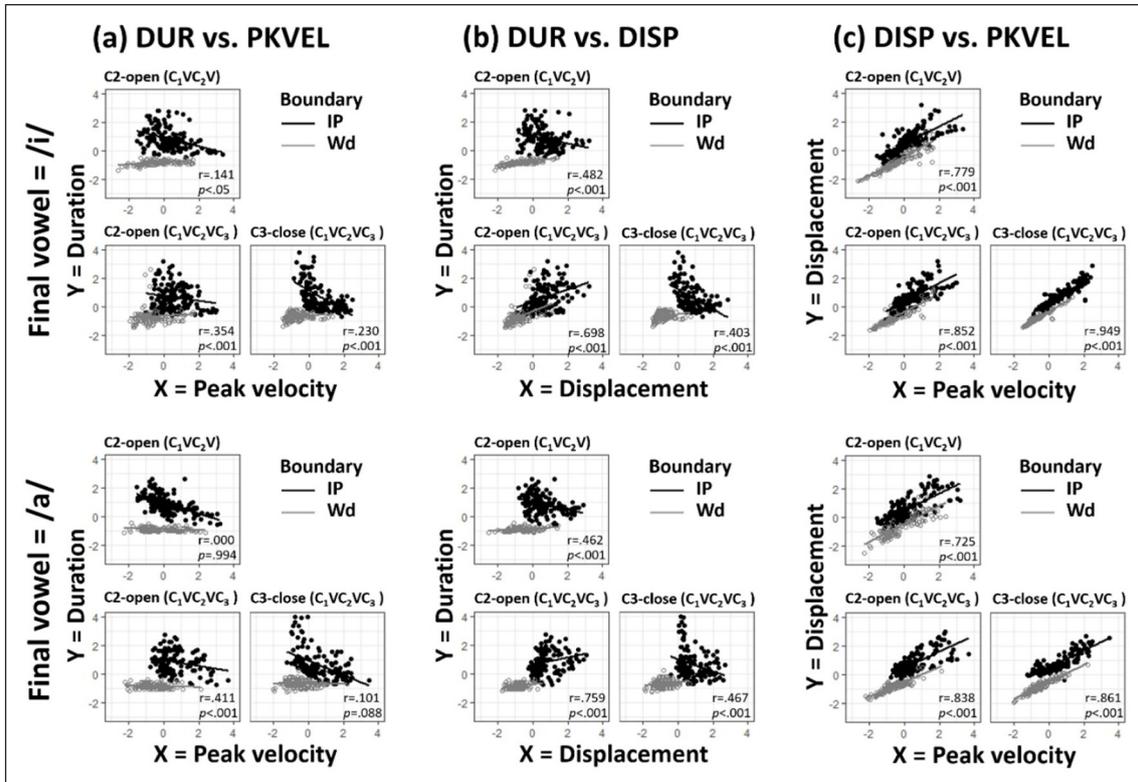
The effect of Info-Structure also did not consistently interact with the boundary effect. In **Figure 3c–d**, the boundary effect on DISP and PKVEL was generally greater when the target was produced on the given condition, as differentiated by the line type (solid = given, dashed = new). However, as marked by the gray ovals, the Boundary  $\times$  Info-Structure interaction effect on DISP and PKVEL was significant only in the final C3-closing gesture in  $C_1V.C_2VC_3$  words. Temporal measures in **Figure 3a–b** also showed significant interaction effects on C2-opening of  $C_1V.C_2VC_3$ , but the direction of the interaction effect was reversed in the following gesture (C3-closing), showing inconsistent interaction effects.

### 3.2. Relationships between kinematic variations

In this section, we provide further data on kinematic characteristics of preboundary articulation by exploring relationships between individual kinematic parameters associated with gestures in the final syllable. Note that the purpose of these additional analyses was to further illuminate how boundary-related temporal variation as reflected in PBL may be related to variation in displacement and movement velocity that is also observed with preboundary articulation. Understanding these relationships between kinematic parameters will also allow us to speculate on the nature of preboundary articulation in dynamical terms. In particular, if variation in spatial expansion that was found to accompany PBL can be largely accounted for by variation in PBL, one can infer that the temporal modulation of gesture at a prosodic juncture is what underlies spatial expansion at that prosodic juncture. Alternatively, if the relationship analyses reveal that variation in spatial expansion is independent from variation in PBL, one can assume that there is an additional dynamical parameter that may be modulated to give rise to variation in spatial expansion.

**Figure 4** visualizes the kinematic relationships in scatter plots with an indication of the boundary condition (IP-final vs. IP-medial) to which each datapoint belongs. However, to examine correlations between the individual parameters within the overall distribution, rather than within each boundary level, the  $r$ -coefficients and  $p$ -values provided in **Figure 4** are based on all data points pooled across boundary conditions. The scatter plots are presented separately for the vowel sequence condition (i.e., whether the final vowel was /i/ or /a/), because the two final vowels are not comparable in terms of the size of oral opening and intrinsic vowel length.

As shown in **Figure 4a**, it is interesting to see that DUR is only weakly correlated with PKVEL ( $r$ -coefficients ranging from 0 to 0.41). Instead, the data points of IP-final are separable from those of IP-medial primarily on the vertical (DUR) dimension, indicating that despite some degree of correlation between the two parameters, the temporal expansion at a larger prosodic boundary (IP-final vs. IP-medial) is largely independent from variation in movement peak velocity. On the other hand, as shown in **Figure 4b**, DUR is quite highly correlated with DISP ( $r$ -coefficients ranging from 0.40 to 0.76), indicating a better correlation between variations in duration and displacement than between variations in duration and movement peak velocity.



**Figure 4:** Relationships between kinematic parameters of DUR, PKVEL, and DISP. Data points are plotted by boundary conditions along with separate linear regression lines (black and gray). (Note that the values on the X and Y axes reflect the mean normalized z-scores obtained from each speaker’s z-scores.)

Accordingly, as can be seen in **Figure 4b**, the distribution of IP-final versus IP-medial data points is separable diagonally (i.e., on both vertical and horizontal dimensions), indicating that PBL in the final syllable cannot be accounted for by variation in DUR or DISP alone, but by their interdependent relationship. Finally, as shown in **Figure 4c**, DISP is highly correlated with PKVEL (the larger, the faster), with r-coefficients ranging from 0.73 to 0.95. Though to a lesser degree than **Figure 4b**, the data points for IP-final are distributed upward to the right relative to those for IP-medial, confirming that the spatial expansion of preboundary (IP-final) articulation accompanies an increase in movement peak velocity.

## 4. General discussion

### 4.1. Kinematic characteristics of preboundary articulation

One of the basic questions that the present study started off with was how preboundary articulation in Korean as an edge-prominence language could be characterized in kinematic terms, and whether and how it would differ from that in English as a head-prominence language

(cf. Jun, 2014). Our results indicate that preboundary articulation at the right edge (in the IP-final position) in Korean is not only longer in duration (preboundary lengthening), but it is also larger in displacement and higher in peak velocity. Preboundary articulation in Korean can therefore be best characterized in kinematic terms as being larger in all directions compared to that in the IP-medial position. Moreover, such an articulatory strengthening effect does not interact with Info-Structure (given/new) at all in CVCV, and for most cases in CVCVC. This indicates that preboundary articulation is modulated in both spatial and temporal dimensions by prosodic constituency (or prosodic phrasing), largely independent of ‘new’ versus ‘given’ conditions of Info-Structure.

This boundary-induced kinematic pattern in Korean is quite different from some of the earlier findings in English, which indicate that preboundary articulation is characterized primarily by temporal expansion (PBL), while spatial expansion and increased movement velocity do not necessarily accompany it (e.g., Beckman & Edwards, 1992; Byrd & Saltzman, 1998; Byrd & Saltzman, 2003; Cho, 2006). Instead, the observed kinematic characteristics of preboundary articulation in Korean appear to be more comparable to those of localized hyperarticulation (prominence/stress-induced articulatory strengthening) in English that can also be characterized by a spatio-temporal expansion (i.e., with articulation under prominence being larger, longer and faster compared to that in the non-prominent context) (Beckman & Edwards, 1992; de Jong, 1995; Fowler, 1995; Cho, 2006; see Mücke & Grice, 2014 for similar effects in German). This holds regardless of the broad focus conditions driven by information structure (‘new’ versus ‘given’). Thus, the preboundary articulation in Korean that reveals spatio-temporal articulatory strengthening effects can be taken to make preboundary articulation *prominent* (or perceptually salient) in a similar way that prominence-induced hyperarticulation effects do in English.

The right-edge effect on hyperarticulation in the spatio-temporal dimension found in Korean bears some broader implications when compared to the realization of contrastive tones at the lexical level. For instance, DiCano, Benn, and García (2021) studied tone realization in Yoloxóchitl Mixtec, an endangered Mexican language with fixed stem-final stress and distinct lexical tones. Their acoustic experiments revealed that speakers of Yoloxóchitl Mixtec not only extended the temporal realization of the final syllable, a form of preboundary lengthening comparable to the IP-final position, but also enhanced the contrast of lexical tones. This enhancement included expanding the tonal  $f_0$  range, resulting in high tone raising and lowering of both low and falling tones. Emerging evidence from the present study and previous studies, including DiCano et al. (2021), on different languages suggests that preboundary temporal slowing, initially rooted in low-level phonetics, can manifest as a locus of hyperarticulation at the right edge of prosodic constituents on a phrase level. This phenomenon extends to both spatial and tonal dimensions, and it may be under speaker control, as discussed by Cho (2016) and DiCano et al. (2021) in the context of maximizing lexical distinction through prominence.

Another basic finding of the present study is that the scope of PBL in Korean was not confined to the proximal gestures in the final syllable, but PBL (as measured by movement duration) was found to extend leftward to the distal consonantal closing gesture (C1-closing) and opening gesture (C1-opening) of the non-final syllable in bisyllabic words. The magnitude of PBL of the two leftmost gestures was, however, drastically attenuated, and time to peak velocity did not show a significant boundary effect in these distal gestures associated with the first syllable. This implies that, while the PBL process operates in a gradient fashion across the two syllables, the effect is asymmetrical and weighted much more strongly on the gestures in the final syllable as has been found with other languages. It is also noticeable that the leftward spreading to the most distal gestures was not observed in displacement and peak velocity as well, indicating that the boundary-induced spatio-temporal expansion is largely localized to those gestures that form the final syllable.

The present study also asked how kinematic characteristics of preboundary articulation under the hypothesized edge-related prominence would be compared with those in the ‘new’ (broad focus) versus ‘given’ (no focus) conditions that are related to prominence associated with information structure that may not be always expressed by phrasing in Korean. Results indicate that the edge-induced strengthening effects on preboundary articulation in Korean are indeed different from those of strengthening that might arise with ‘new’ versus ‘given’ conditions of Info-Structure. Recall that there were significant effects of Info-Structure on kinematic realization primarily in displacement and peak velocity across the board. All the lip closing and opening gestures except for C2-opening gesture in the closed syllable ( $C_1V.C_2VC_3$ ) condition (which showed a marginal effect) are produced with a significant increase in both displacement and peak velocity in the ‘new’ (broad focus) condition compared to the ‘given’ (unfocused) condition. On the other hand, the ‘new’ condition of Info-Structure barely induced temporal expansion—i.e., the ‘new’ condition showed a lengthening effect only on C2-closing gesture in  $C_1V.C_2V$ , and only C3-closing gesture in  $C_1V.C_2VC_3$ , whereas all other gestures showed no lengthening effect of Info-Structure. These results suggest that the articulatory strengthening effect of prominence associated with the ‘new’ (broad focus) condition is primarily reflected on the spatial dimension while the articulatory strengthening effect of prominence associated with the right edge is reflected on both the spatial and temporal dimensions.

#### **4.2. Understanding asymmetrical boundary effects due to syllable structure**

The present study also explored whether the presence or absence of the coda gesture (C3-closing) in the final syllable would influence the distribution of PBL, especially given that any gesture that precedes it can be assumed to be more distal from the boundary in its presence ( $C_1V.C_2VC_3$ ) than in its absence ( $C_1V.C_2V$ ). The results revealed that the magnitude of PBL on the C2-closing gesture was comparable in the two syllable contexts (i.e. 47%-increase when followed by two

gestures (C2-opening and C3-closing) in  $C_1V.C_2VC_3$  versus 44%-increase when followed by one gesture (C2-opening) in  $C_1V.C_2V$ ). This indicates that as far as C2-closing gesture is concerned, the physical distance from the prosodic juncture does not influence the distribution and the magnitude of PBL. Crucially, however, the distribution of PBL was found to differ substantially due to the syllable structure on the gesture(s) that follow the C2-closing gesture. The C2-opening gesture in the absence of the following C3-closing gesture (in the open syllable condition of  $C_1V.C_2V$ ) showed a 208% increase, whereas it showed a 105% increase in the presence of the following C3-closing gesture (in the closed syllable condition of  $C_1V.C_2VC_3$ ). On the other hand, the C3-closing gesture itself in CVCVC context showed a 123% increase.

These results have some implications for effects of syllable structure of the final syllable (open versus closed) on which PBL is concentrated. When we consider the temporal distribution of PBL separately in each syllable context, it still shows a progressively attenuating PBL effect from the right edge to the left. But it is noteworthy that the magnitude of PBL on the C2-opening gesture in the open syllable context ( $C_1V.C_2V$ ) was far greater than that on either the C2-opening gesture or on the C3-closing gesture in the closed syllable ( $C_1V.C_2VC_3$ ). In fact, the magnitude of PBL on the C2-opening gesture (208%-increase) in the open syllable context was largely comparable to a magnitude of PBL on the C2-opening and C3-closing gestures combined (105%-increase and 123%-increase, respectively) in the closed syllable context. Interestingly, a similar effect has been observed in an acoustic study on Japanese PBL (Seo et al., 2019), showing that the magnitude of PBL in the open syllable with one mora is a combined effect of the final two morae (or a vowel and a nasal coda) that form a rhyme of the closed syllable. The observed effects may be due to PBL being realized on the basis of syllable structure with a similar magnitude on the rhyme (as discussed in Seo et al., 2019), or on the basis of the type of the constriction gesture (closing versus opening) that is aligned with the right edge. While this remains to be further corroborated, what appears to be the case is that the magnitude of PBL on the final syllable is influenced by whether there is a coda constriction or not. On a related point, the boundary effect on other kinematic measures (displacement and peak velocity) was found to be more robust on the final gestures when the syllable was closed (when the final gesture involves a closing movement) than when it was open. Again, it is not clear to us why this asymmetry occurred, except that it was due to the presence/absence of the final closing (coda) gesture. It may be due to the fact that the closed syllable is phonologically heavy, so that the boundary effect that lends prominence is more robustly realized on the spatial dimension in Korean. The fact that the broad focus effect was also found to be most robust on the final coda closing gesture appears to lend further support to the heavy-syllable related possibility. (Recall that the prominence arising with the 'new' (broad focus) condition of Info-Structure was primarily reflected on a spatial expansion across the board, but the rightmost C3-closing gesture in the closed syllable condition showed robust effects in both spatial and temporal dimensions.) Alternatively, or in addition, the asymmetry may be due

to the fact that the closed syllable involves the constriction formation gesture for occlusion that is not released in Korean, which may cause more pronounced boundary-related articulation of the formation gesture.

### **4.3. Understanding kinematic characteristics of preboundary articulation of Korean in dynamical terms**

Examination of the relationships between kinematic parameters further revealed the nature of the observed boundary-induced articulatory strengthening in Korean. It was found that PBL was only weakly related to variation in movement velocity, while it was more closely related with spatial expansion as reflected in variation in displacement, which in turn was highly correlated with variation in peak velocity. These results elaborate on the strengthening of preboundary articulation in dynamical terms in Korean in relation to the theory of the  $\pi$ -gesture model (Byrd & Saltzman, 2003). As introduced at the outset of the paper, the  $\pi$ -gesture, as a temporal modulation gesture, is assumed to modulate the rate of the clock in a dynamical system that controls articulatory temporal activation in the vicinity of the prosodic juncture. Thus, at first glance, spatial expansion of the sort observed in this study may not be considered to come about as direct consequence of  $\pi$ -gesture as a temporal modulation gesture. But it is still explainable in the  $\pi$ -gesture model, as clearly noted by Byrd and Saltzman (2003). For example, greater activation time of the  $\pi$ -gesture can provide sufficient time for the gestural target to be attained fully without truncation by an upcoming gesture, which allows possible spatial expansion of preboundary articulation. In this regard, Byrd and Saltzman (2003, p. 172) explain that “[i]t seems that transgestural perturbations of clock rate due to a  $\pi$ -gesture that locally slows time flow in an utterance can result in appropriate kinematic changes not only in the temporal domain but also in the spatial domain.” In fact, in the present study, the time-to-peak velocity measure, which is not affected by possible truncation, indicated a clear PBL effect on the final syllable as predicted by the  $\pi$ -gesture model. One might still ask, however, whether such spatial variation, if it occurs purely due to the absence of gestural truncation, can account for an increase in peak velocity because gestural truncation itself does not directly modify movement velocity. We do not have a compelling explanation to offer, but it may be simply due to the natural tendency that the articulatory movement peak velocity is linearly related to the movement displacement (Munhall et al., 1985; Ostry & Munhall, 1985)—i.e., the increased peak velocity may be attributable to this relationship (see Roon, Hoole, Zeroual, Du, & Gafos, 2021 for related discussion on the relatively little role in variation in peak velocity which explains differences in gestural overlap of consonant clusters in Moroccan Arabic).

Now an important question with respect to language specificity arises from the fact that preboundary articulation in Korean is much stronger in terms of spatio-temporal realization (especially reflected in the spatial expansion) than in English, and that it is by and large comparable

to the prominence-induced hyperarticulation in English. Thus, while preboundary articulation in both English and Korean can be taken to be modulated by the  $\pi$ -gesture as discussed above, we are left with a question regarding how the Korean-specific strengthening pattern associated with preboundary articulation can be understood in the theoretical framework of the  $\pi$ -gesture. Admittedly, this question cannot be easily answered with the limited data available in the present study, but one possibility is that the  $\pi$ -gesture interacts with another kind of modulation gesture in a language-specific way that regulates articulation in the spatial dimension, which in turn engenders hyperarticulation of the sort observed with preboundary articulation in Korean. In this regard, Saltzman, Nam, Krivokapic, and Goldstein (2008) extended the concept of the  $\pi$ -gesture as a temporal modulation gesture to a more general modulation gesture, the  $\mu$ -gesture. They considered two kinds of  $\mu$ -gesture (i.e., a temporal modulation gesture (a “ $\mu_t$ -gesture”) and a spatial modulation gesture (a “ $\mu_s$ -gesture”). They suggested that the prominence-related strengthening pattern (i.e., localized hyperarticulation) can be accounted for by an interaction of two types of the  $\mu$ -gesture that regulate the spatial variation and the temporal variation of the articulation movement of constriction gestures, respectively. The temporal modulation gesture, the ‘ $\mu_t$ -gesture,’ shares similarities with the  $\pi$ -gesture, as both control the temporal aspects of constriction gestures. If the observed spatial expansion, accompanied by an increase in peak velocity, exceeds what can be explained by clock slowing alone under the influence of the temporal modulation of the  $\pi$ -gesture, we should consider the introduction of a spatial modulation gesture of some sort that interacts with the temporal modulation of the  $\pi$ -gesture (see Iskarous & Pouplier, 2022, for a related discussion). Just as a general  $\mu$ -gesture integrates both the spatial and the temporal modulation, it is in principle feasible that the  $\pi$ -gesture can be extended to integrate a spatial modulation gesture all within the theoretical framework of the (‘prosodic’)  $\pi$ -gesture. Such an integration can account for the hyperarticulation effect observed at the right edge in our study, as well as the consistent domain-initial (left-edge) strengthening typically observed in Korean.

But the general progressive pattern of PBL observed in Korean is still predicted by the  $\pi$ -gesture model (Byrd & Saltzman, 2003). As discussed in Cho (2006), activation time can increase and decrease differently across languages, resulting in possible cross-linguistic variation in PBL. The  $\pi$ -gesture activation time in Korean can be considered to be stretched to the extent that it controls temporal realization of the distal gestures in the first (non-final) syllable. Moreover, the activation strength of the  $\pi$ -gesture is assumed to be maximal at the prosodic juncture and then decreases gradually towards the distal gestures in both directions (left and right). This can account for the drastic attenuation of the PBL of the distal gestures. It should be noted, however, that the current finding of the leftward spreading of PBL to the initial syllable is in contrast with a shortening of the penultimate syllable reported in Jang and Katsika (2020). Thus, we should not draw a firm conclusion on this, but we suspect that some differences between the two studies

may have resulted in the discrepancy. First, two-syllable target words used in Jang and Katsika occurred as a second noun in a two-word compound noun phrase. Second, the vowel in the penultimate syllable of the target word which showed a shortening effect was an intrinsically short vowel /i/. Third, their target words occurred always in the non-prominent (unfocused) context with contrastive (narrow) focus falling on the first noun of the compound roughly in the half of their tokens analyzed (in their four-syllable-long Accentual Phrase context), or on an earlier word in the other half of the tokens analyzed (in their seven-syllable-long Accentual Phrase context). Thus, their target words occurred in a context of post-focal reduction and dephrasing. One cannot therefore rule out the possibility that their reported small degree of shortening of the non-final syllables has to do with these specific conditions which all appear to contribute some kind of reduction to the non-final syllable. But given that these specific conditions may affect both the phrase and the word boundary conditions tested in their study, how exactly these differences would lead to discrepant results reported in the present study remains to be further investigated. More research on PBL in various other contexts is certainly called for to understand the exact nature of PBL effects on non-final syllables.

## 5. Conclusion

In conclusion, Korean shows language-specifically fine-tuned kinematic variation of preboundary articulation, seemingly appropriate for an edge-prominence language. PBL is accompanied by spatial expansion, showing articulatory strengthening at the right edge that appears to be much more robust than what has been observed in other languages, such as English. There is some evidence for PBL to spread leftwards to the first syllable of the disyllabic test word, but at an extremely low magnitude, thus leaving the question open as to whether the scope of PBL is set to do so in Korean. PBL also appears to be conditioned by the syllable structure or the gestural content (whether the rightmost gesture is opening versus closing). These results are largely accounted for by the basic concepts of the theory of  $\pi$ -gesture, but it still remains to be seen how language-specifically tuned strengthening of preboundary articulation in Korean, which can be characterized as the constituent gestures being longer, larger and faster, can be modeled in dynamical terms in theories of speech production.

---

## Appendix 1. Model structure and output of the main analysis

### 1. Model structure

Each of the four dependent variables (DV) was modelled separately for each gesture using the following lme4 syntax in R.

`lmer (DV ~ Boundary * Info-Structure + C-type + V-type + Boundary:C-type + Boundary:V-type + (1 + Boundary |Subject) + (1 |TestWord))`

### 2. Model output

$\beta$ -coefficients and  $p$ -values for the intercept and the five fixed effects that are not given in Table 2 in the results section are summarized below in Table 3.

Word type		CV.CV				CV.CVC				
Gesture ID		C1-closing	C1-opening	C2-closing	C2-opening	C1-closing	C1-opening	C2-closing	C2-opening	C3-closing
DV = DUR	Intercept	95.889 $p < .001$	73.770 $p < .001$	101.428 $p < .001$	150.391 $p < .001$	93.269 $p < .001$	72.100 $p < .001$	96.831 $p < .001$	103.956 $p < .001$	117.014 $p < .001$
	C-type	6.139 $p < .001$	-8.234 $p = .078$	31.831 $p = .086$	28.003 $p = .116$	7.163 $p = .961$	-7.101 $p = .853$	26.054 $p = .171$	27.446 $p = .065$	-20.770 $p = .165$
	V-type	-7.506 $p < .001$	-0.665 $p = .630$	1.067 $p = .846$	3.524 $p = .619$	-5.185 $p = .132$	-0.901 $p = .519$	-2.586 $p = .780$	5.282 $p = .310$	0.242 $p = .972$
	Bnd:Info	-0.292 $p = .841$	-0.675 $p = .623$	0.510 $p = .787$	4.127 $p = .453$	-1.784 $p = .256$	-0.436 $p = .722$	-1.062 $p = .662$	-11.659 $p < .001$	15.177 $p = .005$
	Bnd:C-type	6.818 $p < .001$	0.635 $p = .643$	31.407 $p < .001$	25.383 $p < .001$	5.119 $p = .001$	-1.353 $p = .270$	40.947 $p = .106$	47.051 $p < .001$	-30.404 $p < .001$
	Bnd:V-type	0.360 $p = .805$	-3.393 $p = .014$	-6.232 $p < .001$	-7.603 $p = .167$	-1.297 $p = .408$	-2.782 $p = .023$	2.950 $p = .225$	19.435 $p < .001$	-9.790 $p = .069$
DV = T_PV	Intercept	30.814 $p < .001$	40.354 $p = .003$	25.000 $p = .034$	46.299 $p < .001$	30.349 $p < .001$	40.448 $p < .001$	24.400 $p = .042$	43.887 $p < .001$	33.122 $p < .001$
	C-type	-3.780 $p < .001$	-3.382 $p = .597$	-3.507 $p = .467$	11.118 $p = .121$	-2.470 $p = .106$	-3.191 $p = .574$	-3.539 $p = .497$	7.374 $p = .133$	-0.437 $p = .518$
	V-type	-0.759 $p = .061$	2.000 $p = .739$	-2.014 $p = .639$	-2.076 $p = .509$	-0.039 $p = .940$	1.743 $p = .740$	-2.585 $p = .594$	-3.730 $p = .252$	3.895 $p < .001$

(Contd.)

Word type		CV.CV				CV.CVC				
Gesture ID		C1-closing	C1-opening	C2-closing	C2-opening	C1-closing	C1-opening	C2-closing	C2-opening	C3-closing
	Bnd:Info	-0.387 <i>p</i> = .632	-0.028 <i>p</i> = .976	0.449 <i>p</i> = .452	-0.654 <i>p</i> = .699	0.229 <i>p</i> = .759	0.076 <i>p</i> = .945	0.545 <i>p</i> = .318	-4.866 <i>p</i> = .050	0.159 <i>p</i> = .909
	Bnd:C-type	0.685 <i>p</i> = .397	2.250 <i>p</i> = .018	-2.431 <i>p</i> < .001	6.453 <i>p</i> < .001	-0.546 <i>p</i> = .464	1.313 <i>p</i> = .231	-2.922 <i>p</i> < .001	12.001 <i>p</i> < .001	1.222 <i>p</i> = .366
	Bnd:V-type	0.754 <i>p</i> = .351	-0.875 <i>p</i> = .358	-1.667 <i>p</i> = .005	-13.408 <i>p</i> < .001	0.564 <i>p</i> = .450	-0.223 <i>p</i> = .839	-1.636 <i>p</i> = .003	-8.430 <i>p</i> < .001	5.891 <i>p</i> < .001
DV = DISP	Intercept	1212.744 <i>p</i> < .001	876.030 <i>p</i> < .001	902.865 <i>p</i> < .001	1059.712 <i>p</i> < .001	1215.859 <i>p</i> < .001	876.318 <i>p</i> < .001	902.047 <i>p</i> < .001	931.227 <i>p</i> < .001	867.789 <i>p</i> < .001
	C-type	74.275 <i>p</i> < .001	-143.219 <i>p</i> = .359	-93.099 <i>p</i> = .505	207.033 <i>p</i> = .055	78.433 <i>p</i> = .114	-136.517 <i>p</i> = .424	-94.593 <i>p</i> = .563	190.476 <i>p</i> = .076	60.768 <i>p</i> = .388
	V-type	18.067 <i>p</i> = .139	-390.990 <i>p</i> = .145	-391.300 <i>p</i> = .151	338.394 <i>p</i> = .034	5.929 <i>p</i> = .749	-405.060 <i>p</i> = .165	-420.781 <i>p</i> = .170	293.605 <i>p</i> = .050	381.484 <i>p</i> = .071
	Bnd:Info	5.189 <i>p</i> = .832	-29.999 <i>p</i> = .250	-35.347 <i>p</i> = .192	-55.433 <i>p</i> = .075	3.894 <i>p</i> = .879	-0.049 <i>p</i> = .999	0.321 <i>p</i> = .991	-47.388 <i>p</i> = .087	-82.364 <i>p</i> = .001
	Bnd:C-type	0.992 <i>p</i> = .968	44.646 <i>p</i> = .087	-18.719 <i>p</i> = .489	-11.530 <i>p</i> = .711	-5.810 <i>p</i> = .820	19.201 <i>p</i> = .490	-24.898 <i>p</i> = .395	172.432 <i>p</i> < .001	192.186 <i>p</i> < .001
	Bnd:V-type	6.520 <i>p</i> = .789	-20.090 <i>p</i> = .441	-14.928 <i>p</i> = .581	-79.141 <i>p</i> = .011	31.309 <i>p</i> = .221	-0.047 <i>p</i> = .999	19.200 <i>p</i> = .512	82.229 <i>p</i> = .003	-15.634 <i>p</i> = .544
DV = PKVEL	Intercept	283.441 <i>p</i> < .001	190.147 <i>p</i> < .001	217.923 <i>p</i> < .001	183.178 <i>p</i> < .001	287.647 <i>p</i> < .001	193.408 <i>p</i> < .001	219.171 <i>p</i> < .001	169.895 <i>p</i> < .001	181.218 <i>p</i> < .001
	C-type	39.465 <i>p</i> < .001	-22.289 <i>p</i> = .394	-3.028 <i>p</i> = .342	-5.686 <i>p</i> = .677	38.624 <i>p</i> = .047	-19.450 <i>p</i> = .440	-1.865 <i>p</i> = .731	2.062 <i>p</i> = .861	19.271 <i>p</i> = .311
	V-type	2.390 <i>p</i> = .372	-87.075 <i>p</i> = .115	-92.170 <i>p</i> < .001	71.242 <i>p</i> = .091	-2.285 <i>p</i> = .570	-91.380 <i>p</i> = .111	-94.819 <i>p</i> = .028	64.814 <i>p</i> = .091	72.779 <i>p</i> = .089
	Bnd:Info	2.309 <i>p</i> = .666	-5.791 <i>p</i> = .291	-9.088 <i>p</i> = .154	-6.227 <i>p</i> = .287	1.703 <i>p</i> = .759	0.795 <i>p</i> = .888	-2.205 <i>p</i> = .723	-5.477 <i>p</i> = .290	-21.226 <i>p</i> = .001
	Bnd:C-type	-4.969 <i>p</i> = .353	7.233 <i>p</i> = .187	0.682 <i>p</i> = .915	-29.753 <i>p</i> < .001	-0.012 <i>p</i> = .998	3.201 <i>p</i> = .572	-1.700 <i>p</i> = .785	-32.232 <i>p</i> < .001	37.904 <i>p</i> < .001
	Bnd:V-type	-8.013 <i>p</i> = .135	4.061 <i>p</i> = .458	5.182 <i>p</i> = .416	-7.276 <i>p</i> = .214	8.373 <i>p</i> = .131	9.904 <i>p</i> = .081	11.569 <i>p</i> = .064	-0.046 <i>p</i> = .993	-51.260 <i>p</i> < .001

**Table 3:** Summaries of results of LME models. Significant effects are in dark gray.

## Acknowledgements

We would like to thank Dani Byrd on her constructive and insightful comments on an earlier version of the manuscript. In particular, our discussion on how the obtained data could be interpreted in dynamical terms benefited greatly from her inputs. We are also grateful to anonymous reviewers and Associate Editor, Christian DiCanio who provided very constructive and useful feedback on an earlier version of this manuscript on various aspects of the study. Finally, a special thanks goes to Yuna Baek for collecting and organizing the EMA data and measuring kinematics. A preliminary version of this study (with Yuna Baek as a co-author) was presented at the 19<sup>th</sup> International Congress on Phonetic Sciences (ICPhS) 2019. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5C2A02086884) awarded to T. Cho.

## Competing Interests

The authors have no competing interests to declare.

## Author Contributions

Jonny Jungyun Kim is the first author who was responsible for the statistical modeling part and analyses of the data; wrote a substantial portion of the method and results sections; and provided some interpretations of the results. Sahyang Kim participated in the study as the second author who conceptualized the study; provided substantial interpretations of the results; and edited the whole manuscript. Taehong Cho is the corresponding author who supervised and conceptualized the study; provided substantial interpretations of the results; and wrote substantial parts of the introduction and discussion sections.

---

## References

- Barnes, J. A. (2002). *Positional neutralization: a phonologization approach to typological patterns*. Doctoral dissertation, University of California, Berkeley, CA.
- Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, 11(1–2), 17–68. DOI: <https://doi.org/10.1080/016909696387213>
- Beckman, M. E., & Edwards J. (1992). Intonational categories and the articulatory control of duration. In Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka (Eds.), *Speech perception, production, and linguistic structure* (pp. 359–375). Tokyo, Japan: Ohmsha, Ltd.
- Berkovits, R. (1993). Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics*, 21, 479–489. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30231-1](https://doi.org/10.1016/S0095-4470(19)30231-1)
- Berkovits, R. (1994). Durational effects in final lengthening, gapping, and contrastive stress. *Language and Speech*, 37, 237–250. DOI: <https://doi.org/10.1177/002383099403700302>

- Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, 57, 3–16. DOI: <https://doi.org/10.1159/000028456>
- Byrd, D., Kaun, A., Narayanan, A., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: acquisition and the lexicon* (pp. 70–88). Cambridge, UK: Cambridge University Press.
- Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: on the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, 120, 1589–1599. DOI: <https://doi.org/10.1121/1.2217135>
- Byrd, D., & Krivokapić, J. (2021). Cracking prosody in Articulatory Phonology. *Annual Review of Linguistics*, 7, 31–53. DOI: <https://doi.org/10.1146/annurev-linguistics-030920-050033>
- Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, 26, 173–199. DOI: <https://doi.org/10.1006/jpho.1998.0071>
- Byrd, D., & Saltzman, E. (2003). The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149–180. DOI: [https://doi.org/10.1016/S0095-4470\(02\)00085-2](https://doi.org/10.1016/S0095-4470(02)00085-2)
- Cambier-Langeveld, T. (2000). *Temporal marking of accent and boundaries* (Doctoral dissertation). University of Amsterdam. 2000. Available from: <https://www.lotpublications.nl/temporal-marking-of-accent-and-boundaries>
- Cho, T. (2005). Prosodic strengthening and featural enhancement: evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America*, 117, 3867–3878. DOI: <https://doi.org/10.1121/1.1861893>
- Cho, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In L. M. Goldstein, D. H. Whalen & C. T. Best (Eds.), *Laboratory phonology 8: Varieties of phonological competence* (pp. 519–548). Berlin/New York: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110197211.3.519>
- Cho, T. (2015). Language effects on timing at the segmental and suprasegmental levels. In M. Redford (Ed.), *The handbook of speech production* (pp. 505–529). NJ: Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781118584156.ch22>
- Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass*, 10, 120–141. DOI: <https://doi.org/10.1111/lnc3.12178>
- Cho, T., & Jun, S.-A. (2000) Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, 36(CLS 36), 31–44. (An earlier version appeared in *UCLA Working Papers in Phonetics*, 99, 57–70, 2000).
- Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155–190. DOI: <https://doi.org/10.1006/jpho.2001.0131>
- Cho, T. & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4), 466–485. DOI: <https://doi.org/10.1016/j.wocn.2009.08.001>
- 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. DOI: <https://doi.org/10.1017/S0025100315000481>

- de Jong, K. (1995). The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 91, 491–504. DOI: <https://doi.org/10.1121/1.412275>
- DiCanio, C., Benn, J., & García, R. C. (2021). Disentangling the Effects of Position and Utterance-Level Declination on the Production of Complex Tones in Yoloxóchitl Mixtec. *Language and Speech*, 64(3), 515–557. DOI: <https://doi.org/10.1177/0023830920939132>
- Edwards, J. E., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89, 369–382. DOI: <https://doi.org/10.1121/1.400674>
- Fletcher, J. (2010). The prosody of speech: timing and rhythm. In W. J. Hardcastle, J. Laver & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 523–602). Oxford: Blackwell. DOI: <https://doi.org/10.1002/9781444317251.ch15>
- Fowler, C. A. (1995). Acoustic and kinematic correlates of contrastive stress accent in spoken English. In F. Bell-Berti & J. J. Raphael (Eds.), *Producing speech: Contemporary issues: For Katherine Safford Harris* (pp. 355–373). Long Island, New York: AIP Publishing.
- 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). New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511541599.008>
- Gordon, M. (2016). *Phonological Typology*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199669004.001.0001>
- Gussenhoven, C. (2008). Types of focus in English. In C. Lee, M. Gordon & D. Büring (Eds.), *Topic and focus. Studies in linguistics and philosophy* (pp. 83–100). Dordrecht, Netherlands: Springer. DOI: [https://doi.org/10.1007/978-1-4020-4796-1\\_5](https://doi.org/10.1007/978-1-4020-4796-1_5)
- Gussenhoven, G., & Rietveld, A. C. M. (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, 20, 283–303. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30636-9](https://doi.org/10.1016/S0095-4470(19)30636-9)
- Iskarous, K., & Pouplier, M. (2022). Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology. *Journal of Phonetics*, 95(2), 1–28. DOI: <https://doi.org/10.1016/j.wocn.2022.101195>
- Jang, J., & Katsika, A. (2020). The amount and scope of phrase-final lengthening in Korean. In Proceedings of the tenth International Conference on Speech Prosody (pp. 270–274). DOI: <https://doi.org/10.21437/SpeechProsody.2020-55>
- Jeon, H. & Nolan, F., (2017). Prosodic Marking of Narrow Focus in Seoul Korean. *Laboratory Phonology*, 8(1), 2. DOI: <https://doi.org/10.5334/labphon.48>
- Jun, S-A. (1993). *The Phonetics and Phonology of Korean Prosody* (Doctoral dissertation). The Ohio State University, OH.
- Jun, S-A. (2000). K-ToBI (Korean ToBI) labelling conventions. *Speech Sciences*, 7(1), 143–170.
- Jun, S-A. (2014). Prosodic typology: by prominence type, word prosody, and macro-rhythm. In S-A Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 520–540). Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199567300.001.0001>

- Katsika, A. (2016). The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of Phonetics*, 55, 149–181. DOI: <https://doi.org/10.1016/j.wocn.2015.12.003>
- Keating, P. A. (2006). Phonetic encoding of prosodic structure. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 167–186). New York and Hove: Psychology Press.
- Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.), *Papers in laboratory phonology 6: Phonetic interpretations* (pp. 145–163). Cambridge, UK: Cambridge University Press.
- Kentner, G., Franz, I., Knoop, C. A., & Menninghaus, W. (2023). The final lengthening of pre-boundary syllables turns into final shortening as boundary strength levels increase. *Journal of Phonetics*, 97, 101225. DOI: <https://doi.org/10.1016/j.wocn.2023.101225>
- Krivokapić, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, 40, 430–442. DOI: <https://doi.org/10.1016/j.wocn.2012.02.011>
- Krivokapić, J., Styler, W., & Parrell, B. (2020). Pause postures: The relationship between articulation and cognitive processes during pauses. *Journal of Phonetics*, 79, 100953. DOI: <https://doi.org/10.1016/j.wocn.2019.100953>
- Ladd, D. R. (2008). *Intonational Phonology* (2nd ed., Cambridge Studies in Linguistics). UK: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511808814>
- Li, H., Kim, S., & Cho, T. (2023). Preboundary lengthening and its kinematic characteristics in Mandarin Chinese in interaction with focus and lexical tone. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 2149–2153). Guarant International.
- Lindblom, B. (1968). Temporal organization of syllable production. *Speech Transmission Laboratory Quarterly Progress Status Report*, 9, 1–5.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–440). Dordrecht: Kluwer Academic Publishers. DOI: [https://doi.org/10.1007/978-94-009-2037-8\\_16](https://doi.org/10.1007/978-94-009-2037-8_16)
- Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation—Is it mediated by accentuation? *Journal of Phonetics*, 44, 47–61. DOI: <https://doi.org/10.1016/j.wocn.2014.02.003>
- Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 457–474. DOI: <https://doi.org/10.1037/0096-1523.11.4.457>
- Myers, S. (2003). F0 timing in Kinyarwanda. *Phonetica*, 60(2), 71–97. DOI: <https://doi.org/10.1159/000071448>
- Nam, H. (2007). Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In J. Cole & J. I. Hualde (Eds.), *Papers in Laboratory Phonology 9* (pp. 483–506). Berlin: Mouton de Gruyter.
- Ohala, J. J., & Ewan, W. G. (1973). Speed of pitch change. *The Journal of the Acoustical Society of America*, 53(1), 345. DOI: <https://doi.org/10.1121/1.1982441>

- Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, 77, 640–648. DOI: <https://doi.org/10.1121/1.391882>
- Roon, K., Hoole, P., Zeroual, C., Du, S., & Gafos, A., (2021). Stiffness and articulatory overlap in Moroccan Arabic consonant clusters. *Laboratory Phonology*, 12(1), 8. DOI: <https://doi.org/10.5334/labphon.272>
- Saltzman, E., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1, 333–382. DOI: [https://doi.org/10.1207/s15326969eco0104\\_2](https://doi.org/10.1207/s15326969eco0104_2)
- Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira & C. Reis (Eds.), *Proceedings of the fourth International Conference on Speech Prosody* (pp. 175–184). Campinas, Brazil: Capes.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.
- Selkirk, E. (1995). Sentence prosody: intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 550–569). Oxford: Blackwell.
- Seo, J., Kim, S., Kubozono, H., & Cho, T. (2019). Preboundary lengthening in Japanese: To what extent do lexical pitch accent and moraic structure matter? *Journal of the Acoustical Society of America*, 146, 1817–1823. DOI: <https://doi.org/10.1121/1.5122191>
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193–247. DOI: <https://doi.org/10.1007/BF01708572>
- Tiede, M. (2005). *MVIEW: software for visualization and analysis of concurrently recorded movement data*. New Haven, CT: Haskins Laboratories.
- Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35, 445–472. DOI: <https://doi.org/10.1016/j.wocn.2006.12.001>
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91, 1707–1717. DOI: <https://doi.org/10.1121/1.402450>
- Zhang, J. (2014). Tones, Tonal Phonology, and Tone Sandhi. In A. Li & A. Simpson (Eds.), *The handbook of Chinese linguistics* (pp. 443–464). MA: Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781118584552.ch17>

